Ecometrics: A trait-based approach to paleoclimate and paleoenvironmental reconstruction

Vermillion, Wesley

Springer
2018

Vermillion , W , Polly , P D , Head , J , Eronen , J & Lawing , A M 2018 , Ecometrics: A trait-based approach to paleoclimate and paleoenvironmental reconstruction . in D A Croft , S W Simpson & D F Su (eds) , Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities . Vertebrate Paleobiology and Paleoanthropology Series , Springer , pp. 373-394 . https://doi.org/10.1007/978-3-319-94265-0

http://hdl.handle.net/10138/325780
https://doi.org/10.1007/978-3-319-94265-0

unspecified
acceptedVersion

Ecometrics: A Trait-based Approach to Paleoclimate and Paleoenvironmental Reconstruction

Wesley A. Vermillion[1], P. David Polly[2], Jason J. Head[3], Jussi T. Eronen[4,5], A. Michelle Lawing[1*]

[1]Ecosystem Science and Management, Texas A&M University, College Station, TX, USA

[2]Geological Sciences, Biology and Anthropology, Indiana University, Bloomington, IN, USA

[3]Zoology, University of Cambridge, Cambridge, UK CB23EJ

[4]Department of Geosciences and Geography, University of Helsinki and BIOS Research Unit, Helsinki, Finland

* alawing@tamu.edu

Keywords: Functional morphology, Paleoecology, Maximum likelihood, Geographic Distribution, Taxon-free

Suggested running head: Ecometrics: a trait-based approach

ABSTRACT

Ecometrics is a trait-based approach to study ecosystem variability through time. An ecometric value is derived from describing the distribution of functional traits at the community level, which may arise by environmental filtering, extinction, or convergence. An ecometric relationship describes the correspondence between spatially explicit ecometric values and corresponding environmental variation. Transfer functions and maximum likelihood approaches have been developed with modern trait-environment relationships to reconstruct paleotemperature, paleoprecipitation, and paleovegetation cover given the distribution of functional traits within a community. Because the focus for this approach is on the traits and not on species, it is transferable through space and time and can be used to compare no-analog communities. In this paper we review the concepts and history of ecometric analysis and then describe practical methods for implementing an ecometric study.

I. SUMMARY

One of the factors that sort species geographically into communities is the filtering of species by their functional morphologies along climate and environmental gradients (e.g., mammalian herbivores with higher molars eat more gritty and abrasive vegetation and therefore more species with higher molars live in places with more gritty and abrasive vegetation), thus, the morphology has a direct relationship with its environmental condition. In this sorting process, the distribution of functional traits within a community moves toward an optimal environmental condition. These functional traits may be described in terms of their distributions within a community and descriptors of the distribution of community level functional traits have been defined as ecometrics (Eronen *et al.* 2010a; Polly *et al.* 2011). Thus, ecometrics are traits described at the community level that have a relationship with environmental condition. Ecometrics are spatially variable, because the composition of species and traits in communities vary through geography. If species are perfectly sorted, the trait distributions in communities will parallel variation in the environmental condition to which the traits are functionally linked (i.e. the ecometric correlation equals one). If the ecometric correlation is strong, one can estimate the most likely environmental condition given a specific ecometric value using a transfer function or likelihood estimation. If the ecometric correlation is not strong, one may still be able to estimate the most likely environmental condition, but the likelihood surface would be more flat, which would produce a more ambiguous reconstruction. The power of this approach is that functional trait measurements are transferable through space and through time, regardless of the specific species that make up the community. For this reason, ecometrics is a useful approach to reconstruct

paleoclimate and paleoenvironment (Damuth *et al.* 1992).

II. TERMS

Functional trait – a behavioral, biochemical, morphological, phenological, physiological, or structural characteristic that has a physical or chemical interaction with a specific environmental factor. This characteristic is expressed as a phenotype of individual organisms and it is considered relevant to the response of organisms to the surrounding abiotic conditions. A functional trait is something that can be observed and measured.

Geographic sorting – the process by which trait distributions are arranged in geographic space along gradients of environmental conditions.

Taxon free – not depending on taxonomic identity (i.e. it is a descriptor based on trait values rather than on presence or absence of particular taxa). For example, using palms as indicators of frost-free conditions is a taxon-based approach, whereas using counts of stomatal density as indicators of carbon dioxide concentration is a taxon-free approach (Beerling *et al.*, 2011).

Ecometric value – a statistic summarizing the distribution of a functional trait within a community at a single geographic location. Means, variances, proportions, and extremes have been used as descriptors in previous studies; however, any parameter that describes a distribution of functional traits may be used.

Ecometric patterning – geography, or spatial variation, of ecometric values across communities. Perfect patterning represents optimal distributions of populations and species. The lack of patterning could represent a lack of useable covariation, a mismatch between the trait and the environment, or it could result from the inappropriate choice of

environmental correlates.

Ecometric correlation – product-moment correlation between ecometric variation and variation in the environmental condition. High correlations evince a close match between the functional requirements of local communities and their abiotic condition.

Ecometric likelihood – the probability of finding a particular environmental condition given the ecometric value of a community at any single geographic location.

Ecometric load – the amount of mismatch between the distribution of functional traits and the selective optimum for the local environment. High loads may indicate risk to changing environmental conditions. Note that high ecometric loads can be present even with high ecometric correlations if traits are weakly but consistently sorted.

Ecometric optimum – the optimal distribution of an ecometric trait in a particular environment. The optimum is more likely to be a distribution rather than a single trait value, especially in locally heterogeneous environments (e.g., a savannah habitat favors both large and small body sizes).

Transfer function – a function that represents the relationship between the ecometric trait values and a relevant environmental variable.

III. THEORETICAL AND HISTORICAL BACKGROUND

Functional traits are features that organisms use to interact with their environment and with other organisms (Diaz and Cabido 2001; Violle *et al.* 2007). Functional traits of species within a community ultimately mediate the relationship between organisms' morphology and their ecology and link organisms to biotic and abiotic conditions present

105 at particular geographic locations at a particular point in time (Ricklefs and Travis 1980;
106 Chapin 1993; Poff 1997). A certain combination of biotic and abiotic conditions will thus
107 favor traits that maximize an organism's performance, either through natural selection or
108 geographic sorting. Natural selection is an evolutionary process that acts on the variation
109 in a population, where individuals that thrive in their environment produce the most
110 offspring and disproportionately share their traits with the next generation (Darwin and
111 Wallace 1858). Species that move geographically in response to environmental change
112 may follow an environment for which they are already adapted and form new
113 assemblages as a result (Jackson and Overpeck 2000). We are not concerned with the
114 trait values of particular species, but with the distribution of traits within communities
115 (e.g., mean, standard deviation, and range; Damuth *et al.* 1992). Even weak sorting
116 processes can produce trait gradients at the community level, as has been demonstrated in
117 climate gradients in plant leaf traits (Wolfe 1979; Reich *et al.* 1997; Wilf 1997; Wright *et*
118 *al.* 2004; Royer *et al.* 2005; Wing *et al.* 2005; Cornwell and Ackerly 2009; Peppe *et al.*
119 2011), large mammalian herbivores dental traits (Janis and Fortelius 1988; Damuth *et al.*
120 1992; Fortelius *et al.* 2002; Mendoza *et al.* 2005; Damuth and Janis 2011; Eronen *et al.*
121 2010b, 2012), carnivoran locomotor traits (Polly 2010), snake locomotor traits (Lawing *et*
122 *al.* 2012), and mammal life history, physiological, and range traits (Lawing *et al.* 2016).
123 The study of functional traits that have a relationship with climate or environmental
124 gradients at the community level has been termed ecometrics to distinguish it from
125 ecomorphology of particular species or clades (Eronen *et al.* 2010a; Polly *et al.* 2011).
126 **Figure 1** shows a few examples of the morphology of ecometric traits.
127

[insert **Figure 1** near here]

The relationship between the distribution of functional traits in a community and the corresponding environmental factors in the local environment is likely to be complex. Some trait-environment relationships favor a single ecometric optimum in each local environment such that all species tend to have the same trait value. Stomatal density on plant leaves might be expected to have a single optimal value across all members of a community because carbon dioxide concentrations and evapotranspiration conditions are the same for all. However, locally heterogeneous environments are likely to favor a heterogeneous ecometric optimum. For example, savannah environments tend to favor a combination of large body masses (for animals that can travel long distances to procure food in a comparatively resource poor habitat and defend themselves in the open) and small body masses (for species that can subsist on the resources found in a limited area and can find cover in grasses or tree tops). The nature of the ecometric optimum will dictate how it should be measured. An optimum that favors a single trait value can always be adequately represented with the community mean, but an optimum that favors a complex distribution may be better characterized by variance, skewness, or other measure. Temporal variation in environmental conditions may also shape the distribution of the ecometric optimum. For example, Žliobaitė *et al.* (2016) recently showed that recurring extreme conditions (like drought) were better predictors of dental traits than average conditions, indicating that recurring limiting conditions may be important drivers of community functional trait distributions.

Historical Background

Bailey and Sinnott (1915) conducted the first study of community-level functional traits. They showed plant communities that have more leaves with ridges, or non-entire margins, occur in moist temperate climates that experience warm summers and cold winters. Conversely, plant communities that have leaves with no ridges, or entire leaf margins, occur in dry climates including tropic, alpine and arctic regions, moors, steppes, deserts, saline situations, and other physiologically dry climates. They used the proportion of species within a community with a specified phenotype as their metric to describe the community, what we refer to as an ecometric trait value.

Since Bailey and Sinnott (1915) demonstrated the variation in community leaf shape across a temperature gradient, researchers have used leaf shape and other traits to reconstruct paleotemperature. Most notably, Wolfe (1979, 1993) quantified the relationship between leaf margins and temperature, which led to many quantitative paleoclimate reconstructions from leaf assemblages, called leaf-margin analysis. Contemporary to and since Wolfe's work, researchers have been interested in identifying the function of leaf teeth (see Royer and Wilf (2006) for a history and discussion).

The ecometric approach has grown out of the formalization by Damuth *et al.* (1992) of "taxon-free" characterizations of animal communities. These authors were interested in understanding long-term evolutionary and ecological processes. They argued that distributions of traits such as body size, dietary type, locomotor specialization, and predator-to-prey ratios serve as descriptive indices that can be used to compare important aspects of community function in the Paleozoic, Mesozoic, and Cenozoic or between faunal and floral realms in a single time period, because they capture information about productivity, resource availability, community physiognomy, and resource partitioning,

174 comparisons that cannot be made if community function is characterized by the species
175 that dominate it (such as grassland communities). They noted that what makes an
176 approach "taxon-free" is the focus on distributions of traits for inferring
177 paleoenvironment (i.e., the ecometric distributions), not the inferential pathways that lead
178 to the understanding of the trait-environment relationship, which may be founded on
179 taxon-based observations. The fact that many traits have a limited phylogenetic
180 distribution (e.g., teeth are limited to vertebrates, leaf stomata are limited to plants)
181 means that taxon-free approaches, including ecometrics, are not completely free of time,
182 place, and taxon, but comparisons between communities that have no species in common
183 are nevertheless possible. Andrews and Hixson (2014) recently reviewed and critiqued
184 taxon-free approaches.

185 *Ecometric Examples*

186 One of the best-understood ecometrics is the molar crown height of herbivorous
187 large mammals. Different diets vary in the amount of wear they produce: species that eat
188 abrasive or tough foods, or foods of poor nutritive quality, usually have hypsodont teeth,
189 which have a high crown relative to the root. Hypsodonty is a morphology adaptated to
190 high rates of tooth wear (Janis and Fortelius 1988). Proximal factors including
191 abrasiveness of plant material, such as phytolith content and ambient grit and dust, are
192 therefore likely to play a role in selection for tooth crown height (Janis and Fortelius,
193 1988; Strömberg *et al.* , 2013; Fortelius *et al.* , 2014). Tooth crown height correlates with
194 precipitation in the modern world and geologic past (Fortelius *et al.* 2002; Eronen *et al.*
195 2010a,b; Fortelius *et al.* 2014). Eronen *et al.* (2010b) used regression trees to show that
196 there is a strong correlation (65.8%) between crown tooth height and regional

197 precipitation, between diet and precipitation (66.5%), and when diet and tooth crown
198 height are combined (74.2%). Community hypsodonty levels, therefore, have a strong
199 correlation with annual precipitation, which affects both plant tissue properties and the
200 amount of local grit (Fortelius *et al.* 2014) and may be used as a paleoprecipitation proxy
201 that constrain regional details about vegetation patterns and climate models (Fortelius *et*
202 *al.* 2002).

203 Together with other dental traits, hyposodonty can additionally be used to
204 estimate more indirect properties like net primary productivity (Liu *et al.* , 2012) or more
205 proximate ones like production of volcanic ash (Strömberg *et al.* , 2013). Such complex
206 combinations of proximal and distal environmental factors are likely to be associated
207 with all ecometric traits (Polly and Head 2015). Žliobaitė *et al.* (2016) showed that dental
208 traits are closely linked to vegetation greenness (NDVI) in addition to precipitation and
209 temperature. This is reassuring as NPP estimates used in Liu *et al.* (2012) are computed
210 using precipitation as an input and highly dependent on NPP. NDVI is a direct
211 observation of vegetation greenness and is independent of precipitation and temperature
212 measurements. The NDVI depends on climatic conditions and reflects NPP and thus the
213 availability and quality of herbivores' food. Using data from present-day seasonal
214 environments in Africa, Žliobaitė *et al.* (2016) demonstrate that the dental traits show
215 strong correlations to non-availability of preferred plant foods (e.g., during dry seasons or
216 longer dry periods), rather than the properties of average foods consumed. This is the
217 main functional link between climate and herbivore teeth and closely follows the
218 suggestion by Owen-Smith (2002) (see also discussion in Liu *et al.* 2012 and Fortelius *et*
219 *al.* 2014).

Other dental characteristics (such as dental wear patterns, dental structure, and tooth crown complexity) have also been used as estimates of diet in mammals (Fortelius and Solounias 2000; Evans *et al.* 2007; Eronen *et al.* 2010a; Wilson *et al.* 2012; Evans 2013; Saarinen 2014; Saarinen and Karme 2017). Wilson *et al.* (2012) used the dental shape descriptor of orientation patch count (OPC), which is a measure of dental complexity that correlates well with feeding ecology in extant mammals, to determine the paleodiet of the extinct mammalian clade Multituberculata. They were able to identify a shift from a more carnivorous or omnivorous diet among early multituberculates into a more herbivorous diet around 85 million years ago. This shift coincided with the rise of angiosperms. Since vegetation patterns are driven by climatic conditions, these other dental characteristics could potentially be used as climatic proxies. Evans (2013) reviewed dental shape descriptors, which can be used in ecometric studies and will be useful for describing environmental conditions in the past.

Another well-studied trait in both endothermic and ectothermic vertebrates is body size, initiated by the work of Bergmann (1947). Mean body size in non-flying mammals has been shown to increase as temperature decreases, and therefore mean body size increases geographically toward the poles (Rodriquez *et al.* 2008; Eronen *et al.* 2010a). It has also been documented in ectothermic animals that body size is positively correlated with mean ambient temperature (Makarieva *et al.* 2005). This has allowed for the use of ectothermic animals, to be used to estimate paleotemperatures (Head *et al.* 2009; Polly *et al.* 2011; Head *et al.* 2013). Examples of ecometrics in plants, birds, mammals, and snakes are detailed in Table 1, showing the functional trait and the environmental condition for which it is related.

Table 1. Examples of ecometrics in plants, birds, mammals, and snakes. Environmental conditions are mean annual temperature (MAT), annual precipitation (AP), atmospheric carbon dioxide (pCO2), net primary productivity (NPP), trophic position (TP), dietary classification (DC), and vegetation cover (VC).

Group	Functional Trait	Environmental Condition	Citation
Plants	Leaf margin	MAT	Wolf 1979
Plants	Leaf shape	MAT	Wolf 1990; Royer *et al.* 2005; Peppe *et al.* 2011
Plants	Stomata counts	pCO2	Beerling *et al.* 2002
Birds	Body mass	MAT	Meiri and Dayan 2003
Mammals	Body mass	MAT	Damuth *et al.* 1992
Mammals	Tooth morphology	AP; TP; DC; NPP	Evans 2013
Ungulates	Hypsodonty	AP	Fortelius *et al.* 2002
Carnivorans	Limb proportion	VC; MAT; AP	Polly 2010
Snakes	Body length	MAT	Head *et al.* 2009
Snakes	Tail proportion	VC	Lawing *et al.* 2012

Ecometric patterning

Geographic sorting of traits along environmental gradients is one of three processes by which ecometric patterning can emerge, along with evolutionary adaptation and extinction (Polly *et al.* 2015). Environmental gradients are shown to filter species

253 geographically (Belmaker and Jetz 2012). The filtering effects of environmental gradients
254 may sort species geographically by their traits, resulting in local communities with
255 similar trait values that are correlated with local environmental conditions. The ecometric
256 approach was formalized as a way of synthesizing data from the ecological present and
257 geological past to measure responses of ecosystems to climate change in order to better
258 predict the risks posed by global change over the next millennia (Eronen *et al.* 2010a;
259 Polly *et al.* 2011). It has been used to investigate the impacts and risk of extinction
260 (Wilson *et al.* 2012; Polly and Sarwar 2014; Polly and Head 2015) and to understand
261 non-ecological processes (Lawing *et al.* 2016). The relative roles of geographic sorting,
262 local adaptation, and extinction have been investigated by process-based simulation
263 (Polly *et al.* 2015). Ecometrics and its paleontological precursors have been used to
264 estimate paleobiotic conditions using traits (Damuth *et al.* 1992). For example, ecotherm
265 body size (Head *et al.* 2009; Head *et al.* 2013) and leaf shape (Wolfe 1993; Little *et al.*
266 2010) have been used as paleothermometers and tooth crown height, hyposodonty, has
267 been used to estimate paleoaridity (Fortelius *et al.* 2002).
268 The ecometric approach may be used to forecast community responses to
269 anthropogenic climate change (Barnosky *et al.* , 2017). The concept of ecometric load,
270 which is the mismatch between community-level functional trait values and optimal
271 performance, potentially allows us to estimate the vulnerability of a community to
272 climate or environmental change when the optimal performance of the traits along an
273 environmental gradient is known (Polly *et al.* 2015). If changing conditions increase the
274 ecometric load, communities are more likely to become vulnerable.

Methods for using ecometrics to characterize paleoclimate and paleoenvironments have been expanded extensively. For example, Head *et al.* (2009) made a significant push forward on the construction of ecometric models and the application of the transfer function. Instead of solely using correlative models for inference, they used a mechanistic model from the physiology of poikilothermic metabolism along with modern observations of boid snakes and the temperatures they live in to predict paleotemperature at 58-60 Ma (Paleocene) in the neotropics. They found a minimum mean annual temperature of 30-34°C was needed in the Cerrejon Formation in Colombia during this time in order for the large boid, *Titanoboa*, to survive. This is several degrees warmer than the mean annual temperature of 26-27°C found in that area of the world.

The effects of faunal sampling, extinction, and extirpation on ecometric patterns was investigated by Polly and Sarwar (2014) using resampling and rarefaction methods with calcaneum gear ratios. Gear ratios are related to the locomotor style of an animal and, in turn, related to the animal's habitat. Polly and Sarwar (2014) showed that extinction, extirpation, and range change have minimal effects on ecometric correlations when they affect less than a quarter of the species in North American carnivoran communities. If the correlation between gear ratio and an environmental variable, such as vegetation cover, is either high or low, then extinction of more than 25% of species will alter the correlations, but the strongly and weakly correlated environmental variables will still be distinguished. Local extinctions, extirpation, caused a decline in ecometric correlations; however, up to 75% of species could be lost this way before ecometric patterns were completely lost. This result suggested that even in fossil faunal assemblages with a small proportion of the total fauna, the ecometric patterns may still be

298 recoverable if enough independent assemblages are considered. These authors also
299 showed that global extinction and large-scale extirpation have a less predictable effect on
300 ecometric correlations than do local extinctions. How the ecomtetric pattern is affected
301 by global extinciton depends on which organisms go extinct and what their ecometric
302 correlation was. So, if a species, such as a polar bear, is a large contributor to local
303 ecometric means, and if its range is correlated to an extreme of the environmental
304 variable, such as zero vegetation cover, then its extinction could drive correlation up or
305 down dependent on the circumstance.

306 *Ecometric Modeling*

307 Polly *et al.* (2016a) used stochastic modeling to assess the role of population level
308 processes in the formation of ecometric patterns. They systematically varied parameters
309 like local selection intensity, probability of dispersal, probability of extirpation, gene
310 flow, and ancestral trait value to understand the interaction between geographic sorting
311 processes, trait selection, and clade dynamics in producing ecometric patterns. These
312 authors introduced the concept of ecometric load, an analogy with genetic or mutational
313 load, as the mismatch between the distribution of functional traits in a community (or a
314 group of communities) and the selective optimum for those traits in the local environment
315 (Polly *et al.* 2016a). In cases where ecometric load can be measured, it may help predict
316 whether environmental change is likely to make a community more vulnerable to
317 reorganization or extinction. Interestingly, their models frequently produced parallel trait
318 evolution in different clades, rather than the kinds of phylogenetically correlated trait
319 distributions that are observed in many clades. For example, calcaneum gear ratio, which
320 has strong ecometric sorting in the real world, also has a strong phylogenetic correlation

with about 60% of its variance being explained by phylogeny (Polly *et al.* , 2017). The lack of phylogenetic correlation in the models suggested that empirically observed examples of trait-based phylogenetic sorting of species into communities in the modern world results from large-scale turnover in environments.

Ecometrics and ecogeographical rules

Ecogeographical rules often describe patterns that we would classify as ecometric. Ecogeographic rules are hypotheses about how morphological variation changes along environmental or geographic gradients. They may be rules regarding within-species variation or between-species variation. Across species variation has been the most important component of variation to consider for the ecometric approach.

The most notable ecogeographic rule is Bergmann's Rule (Bergmann 1847; Blackburn *et al.* 1999). It states that body mass correlates with temperature and this has been demonstrated both intraspecifically (Ashton 2002; Freckleton *et al.* 2003) and interspecifically (Blackburn and Hawkins 2004; Diniz-Filho *et al.* 2007). Other ecogeographical rules include Allen's rule (length of appendages in endotherms positively correlate with temperature; Allen 1877), Gloger's rule (pigmentation is correlated with humidity within endothermic species; Gloger 1833), Fox's rule (among small Australian mammals during community assembly, it is more likely that species entering a community will represent different functional groups; Fox 1987) and Jordan's rule (there is an inverse relationship between meristic characters and water temperature; Lincoln *et al.* 1982). See Gaston *et al.* (2008) for a review of ecogeographic rules.

IV. IMPLEMENTING AND ECOMETRIC ANALYSIS

In order to document the existence of an ecometric pattern, three types of data are required: geographic ranges of species, abiotic conditions, and functional trait measurement of species. The spatial resolution of the data and the density of sampling of traits and environmental variables within local communities will depend on the scale of the functional relationship and on the question being addressed. Here we focus on ecometric patterns that emerge at regional, continental, or global scales, but the same principles could be applied to patterns that emerge on landscape scales. To simplify large-scale analysis, we recommend measuring trait distributions by sampling the trait once for each species making up the local community and making the assumption that each species has the same trait value everywhere it occurs. Finer scale analyses might benefit from sampling the frequencies of traits among the individuals in a local community to account for variation in abundance of the species making up the community, or from measuring the local values of traits within each community. Because of our choice of scale, our analysis requires information about the geographic range where species are known to live or where they have the potential to live. For our purposes, local community composition includes all the species whose ranges overlap at a sampling point, an assumption that is reasonable for coarse scale analysis, but which might be unreasonable for a finer-scale landscape analysis.

Geographic Range

Geographic range data are often available from field guides and are becoming more readily available online as spatial shapefiles. For example, NatureServe (www.natureserve.org) and International Union for Conservation of Nature host a

website that has an option to download spatial data that depict the ranges of species as polygons (www.iucnredlist.org). The ranges are typically drawn by experts and represent the general areas of occupation of a species. Species occurrences may be used to draw range maps, if no range maps are available, so long as the occurrence data are relatively randomly distributed across the entire environmental range of the species. Once the geographic ranges are obtained, a sampling scheme for measuring the trait-environment relationship needs to be determined. Sampling sites may be systematically or randomly scattered across the extent of the study area. A list of species that occur at each sample site should then be compiled. Some sources of data for geographic ranges include the NatureServe and IUCN for mammals, birds, amphibians, coral and some fishes (www.iucnredlist.org). Plant distribution data for North America and Europe are available from USDA PLANTS database (plants.usda.gov) and from the European Environmental Agency (eea.europa.eu). The Global Biodiversity Information Facility contains species occurrences (longitude and latitude) of both plants and animals (www.gbif.org). A geographic information system (GIS) is useful to work with geographic ranges and other geographic information. Open source systems are available, such as QGIS, DIVA, or even R. Proprietary software is also available; the most popular products are developed through ESRI. Researchers should check with their institutions for GIS services and licensing.

Environmental Data

Environmental factors can be measured directly in the field at sample sites or collected from maps representing the geographic variability of the environmental condition. For every sample site, the environmental variables that are functionally linked

390 to the trait of interest should be sampled. PRISM Climate Group and WorldClim
391 databases provide relatively high-resolution datasets of climate for the globe
392 (www.prism.oregonstate.edu and www.worldclim.org, respectively). Global potential
393 vegetation is available from UW SAGE (www.washington.edu/research/tools/sage/);
394 global historical vegetation coverage is available from Oak Ridge National Laboratories
395 (dx.doi.org/10.3334/ORNLDAAC/419); global digital elevation models and annual
396 productivity are available from NASA (earthobservatory.nasa.gov/) and USGS
397 (nationalmap.gov/elevation.html); global soil distributions are available from
398 International Institute for Applied Systems Analysis (www.iiasa.ac.at/); global land cover
399 is available from the European Space Agency (www.esa.int/); global nitrogen deposition
400 is available from Oak Ridge National Laboratory (www.ornl.gov); global freeze and thaw
401 status is available from the National Snow and Ice Data Center (nsidc.org/); global
402 terrestrial ecosystems are available from World Wildlife Fund (www.worldwildlife.org).
403 In addition to all of these data sources, there are numerous others to find online and in the
404 published literature. One may use any reliable map as a reference for the environmental
405 condition at sample localities.

406 *Functional Traits*

407 Functional traits should be measured or categorized for each species that has the
408 functional trait of interest within each community. Ideally, one would measure the traits
409 from all the species that occur at each sampling location; however, a species value may
410 be used for all of the locations in which that species occurs. Species values for traits may
411 be collected from measuring specimens directly or from obtaining measures from the
412 published literature. Typically only one value for each species is necessary; however, if

413 there is strong geographic variation, the analysis can include species values that are
414 geographically explicit. If the functional trait is variable within species to the extent that
415 it causes performance differences, then individuals should be sorted along the
416 environmental gradient, but if that variation does not cause performance differences with
417 respect to the environmental gradient, then it is reasonable to ignore the within species
418 variation. The functional traits may be continuous or categorical in nature.
419 The equipment necessary for collecting the measurements of functional traits
420 greatly varies depending on the trait of interest. One may collect trait data from published
421 literature, and thus, no equipment is necessary. If measurements are taken directly from
422 specimens, the researcher might want to photograph specimens or use calipers to take
423 direct measurements from specimens. If photographs are taken to investigate specimens,
424 the orientation of the camera and the orientation of the specimen should be consistent for
425 all photographs and photos should include a scale. For classifications or counts, no
426 equipment is necessary, unless the researcher requires magnification. For geometric
427 morphometrics, a camera and scale bar, a 3D scanner, or other digitizing equipment is
428 necessary.
429 *Data Analysis*
430 For each sample site, a list of the relevant species that co-occur there must be
431 assembled. The distribution of functional traits in the community members can then be
432 measured. For the remaining examples, we will use the trait mean to summarize that
433 distribution, but other distributional statistics such as variation, standard deviation, range,
434 or median may be appropriate depending on the expected relationship between the trait
435 and its associated environmental parameter. It is useful to assemble data from all

436 sampling sites into a table with the mean trait values and the associated environmental
437 values in columns and sites in the rows. If there is more than one environmental variable
438 associated with the trait of interest, then in this table each variable requires its own
439 column. Ecometric correlation can easily be calculated from this table as the Pearson's
440 correlation coefficient (r), or a nonparametric equivalent, of the two columns as a
441 measure of correspondence between the environmental parameter and the ecometric trait.
442 If the performance relationship between trait and environment is known, the ecometric
443 load can be calculated as the mean absolute difference between the observed trait means
444 and the expected values for all rows in the table (i.e. the residual between the observed
445 and expected values). Note that ecometric load is largely independent of the ecometric
446 correlation, except insofar that a weak correlation requires at least some mismatch
447 between performance optimum and realized trait mean and thus cannot have a load of
448 zero (Polly *et al.* 2016a). Ecometric patterns with strong correlations can also have strong
449 loads. Establish a transfer function to determine the specific relationship between the two
450 variables (i.e. it estimates function coefficients). If there are fossil localities to estimate
451 past abiotic conditions, then apply the transfer function to the mean of the functional
452 traits measured from the fossils at a single fossil site. Calculate confidence limits for the
453 estimated abiotic condition.

454 Ecometric estimates of paleoenvironment can be made from a regression-based
455 transfer function or from a likelihood distribution of environmental values given an
456 observed ecometric value. If regression-based methods are used, reduced major axis
457 (RMA) regression may be the most appropriate if there is uncertainty in the independent
458 variable. Transfer functions are prediction equations derived from regressing the

459 environmental variable of interest onto the predictive trait value. Different transfer
460 functions are used depending on the type of trait variable and the type of environmental
461 variable. Other considerations include the type of predicted relationship between the trait
462 and environment and the statistical fitting procedures.

463 Transfer functions have been typically calibrated with modern distributions of
464 species and their functional traits. For optimal use of the transfer function, for both
465 reconstructing abiotic condition and tracking trait distributions through time, the
466 functional trait should be either directly fossilizable or able to be estimated from fossil
467 specimens. For example, hypsodonty may be estimated directly from fossil specimens
468 (Fortelius *et al.* 2002) and body size may be accurately estimated from fragmentary
469 skeletal and dental remains (Alexander 1989; MacFadden 1990).

470

471 **V. STRENGTHS OF APPROACH**

472

473 Ecometrics allows us to quantify the dynamic relationships between organisms
474 and their environment. Arguably, the primary strength is that ecometric traits are
475 predictable and transferable through space and time. Since ecometrics uses taxon-free
476 descriptors of community characteristics, it allows for the comparison of community
477 changes in the modern world to those in deep time (Eronen *et al.* 2010a; Polly *et al.*
478 2011). The species that make up the community in the modern world do not need to be
479 the same as those that make up the paleocommunity; the only thing the two communities
480 need to share is the functional trait under study. Because ecometrics focuses on traits that
481 are directly related to environmental conditions, it can be applied to questions on many

482 geographic and temporal scales (Wolfe 1994; Wether *et al.* 1999; Thompson *et al.* 2001;
483 Fortelius *et al.* 2002). This quality allows ecometrics to be applied at local community
484 levels up to global community levels. It also allows for the comparison of ecometric
485 patterning across these scales. In addition, many ecometric traits are easily measured on
486 both animal and plant fossils (Alexander 1989; MacFadden 1990; Fortelius *et al.* 2002;
487 Royer *et al.* 2005; Head *et al.* 2009).
488
489 **VI. BIASES AND SHORTCOMINGS**
490
491 The incomplete nature of the fossil record may influence the quality of data in
492 ecometric studies. However, Polly and Sarwar (2014) showed that even if only 25% of
493 the species of a community are found in a fossil locality, the ecometric patterns still
494 appear. Taphonomic issues with the fossil preservation should also be taken into account,
495 because the functional traits used in ecometric studies need to be measurable or inferred
496 from the fossil remains. It has often not been the case where this is an issue (Alexander
497 1989; MacFadden 1990; Fortelius *et al.* 2002; Royer *et al.* 2005; Head *et al.* 2009; Peppe
498 *et al.* 2011; Wilson *et al.* 2012) and there are various ways to remove the taphonomic
499 influence from the fossil, thus returning it, or a model of it, to its original state (i.e.,
500 retrodeformation methods; Webster and Hughes 1999; Angielczyk and Sheets 2007).
501 Several important assumptions are made with the ecometric approach, including,
502 that the ecometric relationship doesn't change through time and that the full range of
503 morphologies and environmental conditions are represented (i.e. they are not truncated –
504 there are no biased extinctions and only analogous climate conditions). If these are

505 reasonable assumptions for the ecometric trait in question, one may use an ecometric
506 relationship to reconstruct paleoclimate and paleoenvironment, given a couple more
507 assumptions about the fossil localities. The community of fossils within the fossil
508 localities being studied should represent the actual community of species that co-occurred
509 with the depositional environment. In addition, the functional traits should be measurable
510 in the preserved fossils. More work is needed to quantify coefficients describing the
511 relationship between functional traits and environmental conditions to understand if and
512 how they change through time and space.

513 More research needs to be done concerning intraspecific variation. Intraspecific
514 variation, at least in regards to carnivoran calcaneum gear ratios, is high, and it does not
515 follow community level ecometric patterning. Models seem to support the idea that this is
516 possible due to the high level of gene flow within populations, but the overall signal at
517 the community level overshadows the intraspecific variation of the populations.

518

519 **VII. ECOMETRICS EXAMPLE**

520

521 In this section we provide a worked example of how mean annual temperature can
522 be estimated from the distribution of body masses in North American mammals using a
523 likelihood approach. In addition to presenting the method, we also discuss the theory for
524 why ecometric distributions are related to local environment and how our methodology is
525 related to that theory.

526 When the relationship between trait performance and environment is known, the
527 likelihood function can, in theory, be determined *a priori* (Arnold 1983). Performance-

528 environment relationships can sometimes be estimated for traits recovered in fossils using
529 tools such as finite element analysis, hydrodynamic or mechanical principles, or
530 functional morphological experimentation (e.g., Baumiller *et al.* 1991; Rayfield 2007;
531 Wroe 2008; Shino and Suzuki 2011; Stayton 2011; Polly *et al.* 2016b). However, in
532 many cases the performance-environment relationship is unknown, even when the trait-
533 function relationship is well understood. For example, the functional role of hind-limb
534 gear ratio of mammals is understood (it is related through physical lever principles to the
535 efficiency of hind limb extension, and thus to the mechanics of walking, springing,
536 climbing, and running), but its distribution in a particular environment is difficult to
537 predict *a priori* because it is likely to be influenced by a combination of factors such as
538 vegetation physiognomy, terrain ruggedness, substrate, snow cover, and predator-prey
539 interactions (Klein *et al.* 1987; Polly 2008, 2010; Crête and Larivière 2003). In such
540 cases, likelihoods can be estimated directly from empirical data by making the
541 assumption that the distribution of traits in modern communities is close enough to
542 equilibrium to be representative. Polly and Sarwar (2014) found this to be a reasonable
543 assumption for carnivore limb ratios, but care should be taken before generalizing that
544 result to other traits.
545
546 [insert **Figure 2** near here]
547
548 Likelihoods for paleoenvironmental parameters can be estimated empirically by
549 projecting environmental variables into an ecometric space, which is a mathematical
550 space whose axes are statistical descriptors of the distribution of traits in local

551 communities (**Figure 2A**). Ecometric space is in some ways the conceptual opposite of
552 Hutchinson's (1957) niche space, in which a species' niche is the volume it occupies in a
553 multidimensional space whose axes are environmental variables. Translated into
554 ecometric terms, the niche of a species is circumscribed by the performance of the
555 functional traits that relate it to the environmental factors. An ecometric space inverts that
556 relationship by mapping the distribution of environmental parameters on trait-based axes,
557 except that those axes describe the distribution of traits in community assemblages rather
558 than the traits of a single species. **Figure 2A** shows mean annual temperature mapped
559 into an ecometric space whose axes are defined by the distribution of log body mass (kg)
560 of mammals.

561 To create this ecometric space we sampled mammal faunas at 50 km intervals
562 across the entirety of North America. At each sampling point we calculated the mean and
563 standard deviation of the body masses (kg) of the local mammals from the data compiled
564 by Smith *et al.* (2003) and we recorded the local mean annual temperature (MAT) from
565 Hijmans *et al.* (2005).

566 Estimating the likelihood of MAT from community trait characteristics is then a
567 three-step process. First, the trait data should be binned so that a frequency distribution of
568 the environmental variable can be calculated. The number of bins is arbitrary, but it
569 should be coarse enough to encompass a reasonable number of communities in the
570 densest areas of the trait space, yet fine enough to be biologically meaningful. The ideal
571 bin size would be about as wide as the standard errors of the means and standard
572 deviations in the bin. We divided our ecometric space into 625 bins (25 X 25).

573 Second, the MAT observations in each trait bin were themselves binned to
574 tabulate their frequency distribution. The number of temperature bins is also arbitrary (we
575 used a bin size of 1°C). Once binned, a discrete probability density function (PDF) can be
576 estimated by dividing the bin counts by the total number of observations. A continuous
577 PDF can also be fit to the frequency data. **Figure 2B** shows a continuous PDF for
578 temperature in communities with a mean ln body mass between 3.05 (21.11 kg) and 3.13
579 (22.87 kg) and a standard deviation between 1.05 and 1.10 (black box in **Figure 2A**)
580 based on a Gaussian kernel density estimate with bandwidth of 1°C.
581 With these pieces in place, a maximum likelihood estimate of the environmental variable
582 (MAT) can be made by measuring the trait (body mass) mean and standard deviation at a
583 site of interest and then finding the environmental value that maximizes the PDF. The
584 maximum likelihood estimates of MAT for all the sampling locations in North America
585 are shown in **Figure 2C**, compared with the real MAT values in **Figure 2D**. This
586 estimate (MAT at sampling locations) is reasonable, but tends to be too low in high
587 elevation areas and too warm at lower elevations (see anomaly map in **Figure 2E**). On
588 average, the discrepancy between the real MAT and the estimate based on mammal body
589 masses was 3.9°C, which serves as an approximate standard error (but one that is
590 probably underestimated because the likelihood functions were estimated from the same
591 data on which they were tested).
592 One of the primary advantages of using likelihood and ecometric spaces is that the
593 likelihoods of alternative reconstructions can be compared. For example, the maximum
594 likelihood estimate of MAT for the Central American location highlighted by the black
595 circles in **Figure 2C-E** is 24.95°C with $L(24.95)=0.274$ derived from the PDF for that

location (**Figure 2B**). The real MAT is 25.3°C, which has $L(25.3)=0.265$, which is clearly much more plausible with a likelihood ratio of 0.967 than an MAT of 20°C, where $L(20.0)=0.024$ produces a likelihood ratio of only 0.088. This approach can be developed into a formal likelihood ratio test of alternative hypotheses about paleoenvironment, and it can be combined with prior probabilities from independent paleoenvironmental proxies into a Bayesian framework (Polly and Head 2015).

The second advantage of the likelihood approach is that incommensurable traits can be combined. Mammalian body mass distributions were fairly good predictors of MAT, as might be expected by extrapolating Bergmann's rule, which postulates that subspecies in colder climates will have larger body mass than their conspecifics elsewhere (Bergmann 1947; Scholander 1955; Mayr 1963), to the level of species and communities. However, homeothermic mammals are well-insulated from climate and their body size may not be as good predictors of MAT as that of ectotherms (Head *et al.* 2009). Because large ectothermic animals require warm ambient temperatures to sustain their growth rate, the size range of species in hot climates is expected to be greater than in cold climates (Makarieva *et al.* 2005). Using a dataset of ventral scale count as a proxy for size assembled by Lawing *et al.* (2012), we repeated the likelihood estimation for MAT in North America using the range of log scale count and its standard deviation as the dimensions of the ecometric space (**Figure 2F**). These two parameters are both measures of dispersion and are therefore correlated, yet pick out different aspects of the distribution of body size in local communities. These data yield a similar likelihood function for the site in Central America, but with a narrower peak of predicted MAT (**Figure 2G**).

Because the likelihood functions are probabilities, they can be combined by multiplying

619 them and renormalizing so that the area under them is 1.0. The resulting function gives
620 the likelihood of MAT given the distribution of body masses in mammals and the
621 distribution of ventral scale counts in snakes. This combined likelihood function provides
622 a better estimate of temperature than either trait does alone (**Figure 2H-J**). The mean
623 anomaly between estimated and real temperature based on the combined estimate is only
624 2.4°C, much better than with mammals alone. Sacrificed is the ability to estimate MAT in
625 the northern part of the continent where the climate is too cold for snakes to live.
626 A fully worked ecometrics example and R code is documented in Appendix I.
627

628 **VIII. FUTURE PROSPECTS**

629

630 Several aspects of ecometrics need to be addressed with future research. These
631 areas include—intraspecific variation, abundance, phylogenetics, combining multiple
632 ecometrics, conservation, and ecosystem services. Polly *et al.* (2016a) took first steps to
633 identify and define a theory of ecogeography- moving from pattern to process with
634 ecometric modeling. They used modeling to simulate the affects of changing
635 microevolutionary processes on the ecometric patterning seen at the community level. As
636 this area of interest progresses, eventually we hope to be able to observe ecometric
637 patterns and determine which microevolutionary and macroevolutionary processes caused
638 the patterns we see. One area of ecometric study that deserves further study is the affects
639 of intraspecific variation in ecometric patterns. Polly *et al.* (per comm) show that while
640 intraspecific variation, at least in regards to carnivoran calcaneum gear ratios, is high, it
641 does not follow community level ecometric patterning. Models seem to support the idea

642 that this is possible due to the high level of gene flow within populations, but the overall
643 signal at the community level overshadows the intraspecific variation of the populations.
644 Another future development in the field of ecometrics is the prospect of marine
645 ecometrics (Wallin 1991; Yasuhara *et al.* 2015). Due to the depositional environments
646 found in marine systems, marine fossils are the majority of fossils found (Sepkoski
647 1978). Marine fossils are often used in isotopic studies to reconstruct paleoecology
648 (Bowen 1964; Anderson and Arthur 1983; Krantz *et al.* 1987; Wefer and Berger 1991;
649 Rodrigues *et al.* 2014; Huyghe *et al.* 2015; Reich *et al.* 2015). The changes in isotopic
650 fractionation may follow geographical gradients similar to an ecometric (Marchais *et al.*
651 2013; Lynch-Stieglitz *et al.* 2015; Mackenzie *et al.* 2014; Beard *et al.* 2015). However,
652 the isotopic fractionation process is not a morophological trait, it is a metabolic functional
653 trait, and as such, follows different biological, physical, and chemical rules.
654 New advances in remote sensing capabilities could alter the accuracy of our
655 estimates of both species geographic ranges and the environmental condition in which
656 they occur. Remote sensing appears like it will be quite useful when studying changes in
657 plant community structures over large geographic regions. It allows for a level of detail in
658 which single plants are recognizable across regional scales (Jones and Vaughan 2010).
659 Geometric morphometrics will also allow for the study of ecometrics when a
660 particular functional trait changes size across age, but doesn't change shape (or vice
661 versa). New advances will also allow for 3D modeling of functional traits. This may
662 allow for the identification of aspects of the morphology that are better ecometrics than
663 those being used. It can also allow for better ecometrics since we will no longer be using
664 linear or volumetric measurements, but instead use the full 3D shape. Recent work by

665 Polly (2010) and Wilson *et al.* (2012) has shown that this is a viable avenue of future
666 inquiry.
667
668 **IX. REFERENCES**
669

670 Alexander, R. M. (1989). Optimization and gaits in the locomotion of vertebrates.
671 *Physiological Reviews,* 69, 1199-1227.
672 Anderson, T.F. & Arthur, M. A. (1983). Stable isotopes of oxygen and carbon and
673 their application to sedimentologic and paleoenvironmental problems. In M.
674 A. Arthur, T. F. Anderson, I. R. Kaplan, J. Veizer, & L. S. Land (Eds.),
675 *Stable Isotopes in Sedimentary Geology* (pp. 1 - 151), Tulsa, OK: SEPM Short
676 Course.
677 Andrews, P. & Hixson, S. (2014). Taxon-free methods of palaeoecology. *Annales*
678 *Zoologici Fennici,* 51, 269-284.
679 Angielczyk, K.D. & Sheets, H.D. (2007). Investigation of simulated tectonic
680 deformation in fossils using geometric morphometrics. *Paleobiology,* 33, 125-
681 148.
682 Arnold, S. J. (1983). Morphology, performance and fitness. *American Zoologist,* 23,
683 347-361.
684 Ashton, K. G. (2002). Do amphibians follow Bergmann's rule? *Canadian Journal of*
685 *Zoology,* 80, 708-716.
686 Bailey, I. W. & Sinnott, E. W. (1915). A botanical index of Cretaceous and Tertiary
687 climates. *Science,* 41, 831–834.

688 Barnosky, A. D., E. A. Hadly, P. Gonzalez, J. J. Head, P. D. Polly, A. M. Lawing, J.
689 T. Eronen, D. D. Ackerly, K. Alex, E. Biber, J. Blois, J. Brashares, G.
690 Ceballos, E. Davis, G. P. Dietl, R. Dirzo, H. Doremus, M. Fortelius, H. W.
691 Greene, J. Hellmann, T. Hickler, S. T. Jackson, M. Kemp, P. L. Koch, C.
692 Kremen, E. L. Lindsey, C. Looy, C. R. Marshall, C. Mendenhall, A.S Mulch,
693 A. M. Mychajliw, C. Nowak, U. Ramakrishnan, J. Schnitzler, K. D. Shrestha,
694 K. Solari, L. Stegner, M. A. Stegner, N. Chr. Stenseth, M. H. Wake. Z. Zhang.
695 2017. Merging paleobiology with conservation biology to guide the future of
696 terrestrial ecosystems. *Science*, 355 (6325). doi: 10.1126/science.aah4787
697 Baumiller, T. K., LaBarbera, M. & Woodley, J. D. (1991). Ecology and functional
698 morphology of the isocrinid *Cenocrinus asterius* (Linnaeus) (Echinodermata:
699 Crinoidea): in situ and laboratory experiments and observations. *Bulletin of*
700 *Marine Science*, 48, 731-748.
701 Beard, J. A., Ivany, L. C. & Runnegar, B. (2015). Gradients in seasonality and
702 seawater oxygen isotopic composition along the early Permian Gondwanan
703 coast, *SE Australia: Earth and Planetary Science Letters*, 425, 219-231.
704 Beerling, D., Lomax, B., Royer, D., Upchurch, G. & Kump, L. (2002). An
705 atmospheric pCO2 reconstruction across the Cretaceous-Tertiary boundary
706 from leaf megafossils. *Proceedings of the National Academy of Sciences of*
707 *the United States of America*, 99, 7836–7840.
708 Beerling, D. J., Fox, A., Stevenson, D. S. & Valdes, P. J. (2011). Enhanced
709 chemistry-climate feedbacks in past greenhouse worlds. *Proceedings of the*
710 *National Academy of Sciences*, 108, 9770-9775.

711 Belmaker, J. & Jetz, W. (2012). Regional pools and environmental controls of
712 vertebrate assemblages. *American Naturalist*, 179, 512-523.
713 Bergmann, C. (1847). Ueber die Verhältnisse der Wärmeökonomie der Thiere zu
714 ihrer Grösse. *Gottinger studien*, 3, 595-708.
715 Blackburn, T. M., Gaston, K. J. & Loder, N. (1999). Geographic gradients in body
716 size: a clarification of Bergmann's rule. *Diversity and Distributions*, 5, 165-
717 174.
718 Blackburn, T. M. & Hawkins, B. A. (2004). Bergmann's rule and the mammal fauna
719 of northern North America. *Ecography*, 27, 715-724.
720 Bowen, R. (1964). Oxygen isotope paleotemperature measurements on Mesozoic
721 Belemnoidea and their importance in paleoclimatic studies. In U. Colombo &
722 G. D. Hobson (Eds.) *Advances in organic Geochemistry* (pp. 271-283). New
723 York, NY: Pergamon.
724 Chapin, F. S. III. (1993). Functional role of growth forms in ecosystem and global
725 processes. In J. R. Ehleringer & C. B. Field (Eds.) *Scaling Physiological*
726 *Processes: Leaf to Globe* (pp. 287-312). San Diego, CA: Academic Press.
727 Cornwell, W. K. & Ackerly, D. D. (2009). Community assembly and shifts in plant
728 trait distributions across an environmental gradient in coastal California.
729 *Ecological Monographs*, 79, 109-126.
730 Crête, M. & Larivìere, S. (2003). Estimating the costs of locomotion in snow for
731 coyotes. *Canadian Journal of Zoology*, 81, 1808-1814.
732 Damuth, J. D., Jablonski, D., Harris, R. M., Potts, R., Stucky, R. K., Sues, H. D. &
733 Weishampel, D. B. (1992). Taxon-free characterization of animal

communities. In A. K. Beherensmeyer, J. D. Damuth, W. A. diMichele, R. Potts, H. D. Sues, & S. L. Wing (Eds.) *Terrestrial Ecosystems Through Time: Evolutionary Paleoecology of Terrestrial Plants and Animals* (pp. 183-203). Chicago: University of Chicago Press.

Damuth, J. D. & Janis, C. M. (2011). On the relationship between hyposodonty and feeding ecology in ungulate mammals and its utility in paleaeoecology. *Biological Reviews*, 86, 733-758.

Darwin, C. R. & Wallace, A. R. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Journal of the Proceedings of the Linnean Society of London. Zoology*, 3, 45-50.

Díaz, S. & Cabido, M. (2001). Vive la différence: plant functional diversity matters to ecosystem processes. *Trends in Ecology & Evolution*, 16, 646–655.

Diniz-Filho, J.A.F., Bini, L. M., Rodriguez, M. A., Rangel, T. F. L. & Hawkins, B. A. (2007). Seeing the forest for the trees: partitioning ecological and phylogenetic components of Bergmann's rule in European Carnivora. *Ecography*, 30, 598-608.

Eronen, J. T., Polly, P. D., Fred, M., Damuth, J., Frank, D. C., Mosbrugger, V., Scheidegger, C., Stenseth, N. C. & Fortelius, M. (2010a). Ecometrics: the traits that bind the past and present together. *Integrative Zoology*, 5, 88-101.

Eronen, J., Puolamaki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C. & Fortelius, M. (2010b). Precipitation and large herbivorous mammals I: estimates from present-day communities. *Evolutionary Ecology Research*, 12, 217-233.

757 Eronen, J. T., Fortelius, M., Micheels, A., Portmann, F. T., Puolamäki, K. & Janis, C.
758 M. (2012). Neogene Aridification of the Northern Hemisphere. *Geology*, 40,
759 823-826.
760 Evans, A. R. (2013). Shape descriptors as ecometrics in dental ecology. *Hystrix, the*
761 *Italian Journal of Mammalogy*, 24, 133-140.
762 Fortelius, M., Eronen, J. T., Jernvall, J., Liu, L. P., Pushkina, D., Rinne, J., Tesakov,
763 A. *et al.* (2002). Fossil mammals resolve regional patterns of Eurasian climate
764 change over 20 million years. *Evolutionary Ecology Research*, 4, 1005-1016.
765 Fortelius, M., Eronen, J. T., Kaya, F., Tang, H., Raia, P. & Puolamäki, K. (2014).
766 Evolution of Neogene mammals in Eurasia: environmental forcing and biotic
767 interactions. *Annual Review of Earth and Planetary Sciences*, 42, 579-604.
768 Fox, B. J. (1987). Species assembly and the evolution of community structure.
769 *Evolutionary Ecology*, 1, 201-213.
770 Freckleton, R. P., Harvey, P. H. & Pagel, M. (2003). Bergmann's rule and body size
771 in mammals. *The American Naturalist*, 161, 821-825.
772 Gaston, K.J., Chown, S.L. & Evans, K.L. (2008). Ecogeographical rules: elements of
773 a synthesis. *Journal of Biogeography*, 35, 483-500.
774 Head, J. J., Bloch, J. I., Hastings, A. K., Bourque, J. R., Cadena, E. A., Herrera, F. A.,
775 Polly, P.D. & Jaramillo, C. A. (2009). Giant boid snake from the Palaeocene
776 neotropics reveals hotter past equatorial temperatures. *Nature*, 457, 715-717.
777 Head, J. J., G. F. Gunnell, P. A. Holroyd, J. H. Hutchison, & R. L. Ciochon. (2013).
778 Giant lizards occupied herbivorous mammalian ecospace during the

779 Paleogene greenhouse in Southeast Asia. *Proceedings of the Royal Society of*
780 *London, Series B*., 280, 20130665.
781 Hijmans, R. J., Cameron, J. L., Parra, P. G., Jones, P. G. & Jarvis, A. (2005). Very
782 high resolution interpolated climate surfaces for global land areas.
783 *International Journal of Climatology*, 25, 1965-1978.
784 Huyghe, D., Lartaud, F., Emmanuel, L., Merle, D. & Renard, M. (2015). Palaeogene
785 climate evolution in the Paris Basin from oxygen stable isotope (δ18O)
786 compositions of marine molluscs. *Journal of the Geological Society*, 172,
787 576-587.
788 Jackson, S. T. & Overpeck, J. T. (2000). Responses of plant populations and
789 communities to environmental changes of the late Quaternary.
790 *Paleobiology*, 26, 194-220.
791 Janis, C. M. & Fortelius, M. (1988). On the means whereby mammals achieve
792 increased functional durability of their dentitions, with special reference to
793 limiting factors. *Biological Reviews of the Cambridge Philosophical Society*,
794 63, 197-230.
795 Jones, H. G. & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles,*
796 *techniques, and applications*. New York, NY: Oxford University Press.
797 Klein, D. R., Meldgaard, M. & Fancy, S. G. (1987). Factors determining leg length in
798 *Rangifer tarandus*. *Journal of Mammalogy*, 68, 642-655.
799 Krantz, D. E., Williams, D. F. & Jones, D. S. (1987). Ecological and
800 paleoenvironmental information using stable isotope profiles from living and

fossil molluscs. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 58, 249-266.

Lawing, A. M., Eronen, J. T., Blois, J. L., Graham, C. H. & Polly, P. D. (2016). Community functional trait composition at the continental scale: the effects of non-ecological processes. *Ecography*, EarlyView DOI: 10.1111/ecog.01986

Lawing, A. M., Head, J. J. & Polly, P. D. (2012). The ecology of morphology: the ecometrics of locomotion and macroenvironment in North American snakes. In J. Louys (Eds). *Paleontology in Ecology and Conservation* (pp. 117-146). Berlin: Springer Berlin Heidelberg.

Lincoln, R. J. B. & Clark, G. A. (1982). *A dictionary of ecology, evolution and systematics*. Cambridge, UK: Cambridge University Press.

Little, S. A., Kembel, S. W. & Wilf, P. (2010). Paleotemperature Proxies from Leaf Fossils Reinterpreted in Light of Evolutionary History. *PLoS One*, 5, e15161.

Liu, L., Puolamäki, K., Eronen, J.T., Ataabadi, M.M., Hernesniemi, E. & Fortelius, M., 2012. Dental functional traits of mammals resolve productivity in terrestrial ecosystems past and present. *Proceedings of the Royal Society of London B: Biological Sciences*, 20120211.

Lynch-Stieglitz, J., Polissar, P. J., Jacobel, A. W., Hovan, S. A., Pockalny, R. A., Lyle, M., Murray, R. W., Ravelo, A. C., Bova, S. C., Dunlea, A. G., Ford, H. L., Hertzberg, J. E., Wertman, C. A., Maloney, A. E., Shackford, J. K., Wejnert, K. & Xie, R. C. (2015). Glacial-interglacial changes in central tropical Pacific surface seawater property gradients: *Paleoceanography*, 30, 423-438.

MacFadden, B. J. (1990). *Body size in mammalian paleobiology: estimation and biological implications*. Cambridge, UK: Cambridge University Press.

Makarieva, A. M., Gorshkov, V. G. & Li, B. L. (2005). Temperature-associated upper limits to body size in terrestrial poikilotherms. *Oikos*, 111, 425-436.

MacKenzie, K. M., Longmore, C., Preece, C., Lucas, C. H., & Trueman, C. N. (2014). Testing the long-term stability of marine isoscapes in shelf seas using jellyfish tissues. *Biogeochemistry*, 121, 441-454.

Marchais, V., Schaal, G., Grall, J., Lorrain, A., Nerot, C., Richard, P. & Chauvaud, L. (2013). Spatial Variability of Stable Isotope Ratios in Oysters (*Crassostrea gigas*) and Primary Producers Along an Estuarine Gradient (Bay of Brest, France). *Estuaries and Coasts*, 36, 808-819.

Mayr, E. (1963). *Animal Species and Evolution*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.

Mendoza, M., Janis, C. M. & Palmqvist, P. 2005. Ecological patterns in the trophic-size structure of large mammal communities: a 'taxon-free' characterization. *Evolutionary Ecology Research*, 7, 505-530.

Owen-Smith, R. N. (2002). *Adaptive herbivore ecology: from resources to populations in variable environments*. Cambridge University Press.

Peppe, D. J., Royer, D. L., Cariglino, B., Oliver, S. Y., Newman, S., Leight, E., Enikolopov, G., Fernandez- Burgos, M., Herrera, F., Adams, J. M. & Correa, E. (2011). Sensitivity of leaf size and shape to climate: global patterns and paleoclimatic applications. *New Phytologist*, 190, pp.724-739.

846 Poff, N. L. (1997). Landscape filters and species traits: towards mechanistic
847 understanding and prediction in stream ecology. *Journal of the North*
848 *American Benthological Society*, 16, 391-409.
849 Polly, P. D. (2008). Adaptive Zones and the Pinniped Ankle: A 3D Quantitative
850 Analysis of Carnivoran Tarsal Evolution. In E. Sargis & M. Dagosto (Eds.),
851 *Mammalian Evolutionary Morphology: A Tribute to Frederick S. Szalay* (pp.
852 165-194). Dordrecht, The Netherlands: Springer.
853 Polly, P. D. (2010). Tiptoeing through the trophics: geographic variation in
854 carnivoran locomotor ecomorphology in relation to environment. In A.
855 Goswami & A. Friscia (Eds.), *Carnivoran evolution: new views on phylogeny,*
856 *form, and function* (pp. 347-410). Cambridge, UK: Cambridge University
857 Press.
858 Polly, P. D., Eronen, J. T., Fred, M., Dietl, G. P., Mosbrugger, V., Scheidegger, C.,
859 Frank, D. C., Damuth, J., Stenseth, N. C. & Fortelius, M. (2011). History
860 matters: ecometrics and integrative climate change biology. *Proceedings of*
861 *the Royal Society of London B: Biological Sciences*, rspb20102233.
862 Polly, P. D. & Head, J. J. (2015). Measuring Earth-life transitions: ecometric analysis
863 of functional traits from late Cenozoic vertebrates. In P. D. Polly, J. J. Head,
864 & D. L. Fox (Eds.), *Earth-Life Transitions: Paleobiology in the Context of*
865 *Earth System Evolution* (pp. 21-46). The Paleontological Society Papers 21.
866 New Haven, CT: Yale Press.

867 Polly, P. D & Sarwar, S. (2014). Extinction, extirpation, and exotics: effects on the
868 correlation between traits and environment at the continental level. *Annales*
869 *Zoologici Fennici*, 51, 209-226.
870 Polly, P. D., Lawing, A. M., Eronen, J. T. & Schnitzler, J. (2016a). Processes of
871 ecometric patterning: modeling functional traits, environments, and clade
872 dynamics in deep time. *Biological Journal of the Linnaean Society*, 118, 39-
873 63.
874 Polly, P. D., Stayton, C. T., Dumont, E. R., Pierce, S. E., Rayfield, E. J. &
875 Angielczyk, K. (2016b). Combining geometric morphometrics and finite
876 element analysis with evolutionary modeling: towards a synthesis. *Journal of*
877 *Vertebrate Paleontology*, e1111225, 1-29.
878 Polly, P. D., Fuentes-Gonzales, J., Lawing, A. M., Bormet, A. K., & Dundas, R. G.
879 2017. Clade sorting has a greater effect than local adaptation on ecometric
880 patterns in Carnivora. *Evolutionary Ecology Research*, 18, 187-200.Rayfield,
881 E. J. (2007). Finite element analysis and understanding the biomechanics and
882 evolution of living and fossil organisms. *Annual Review of Earth and*
883 *Planetary Sciences*, 35, 541–576.
884 Reich, P. B., Walters, M. B. & Ellsworth, D. S. (1997). From tropics to tundra: global
885 convergence in plant functioning. *Proceedings of the National Academy of*
886 *Sciences USA*, 94, 13730-13734.
887 Reich, S., Warter, V., Wesselingh, F. P., Zwaan, J. C., Lourens, L. & Renema, W.
888 (2015). Paleoecological significance of stable isotope ratios in Miocene
889 tropical shallow marine habitats (Indonesia): *Palaios*, 30, 53-65.

890 Ricklefs, R. E. & Travis, J. (1980). A morphological approach to the study of Avian
891 community organization. *The Auk*, 97, 321-338.
892 Rodrigues, G. B., Fauth, G., Santos, R. V., Koutsoukos, E. A. M. & Colin, J.-P.,
893 (2014). Tracking paleoecological and isotopic changes through the K-Pg
894 boundary from marine ostracodes: The Poty quarry section, northeastern
895 Brazil. *Cretaceous Research*, 47, 105-116.
896 Rodríguez, M.Á., Olalla- Tárraga, M.Á. & Hawkins, B.A. (2008). Bergmann's rule
897 and the geography of mammal body size in the Western Hemisphere. *Global*
898 *Ecology and Biogeography*, 17, 274-283.
899 Royer, D.L. & Wilf, P. 2006. Why do toothed leaves correlate with cold climates?
900 Gas exchange at leaf margins provides new insights into a classic
901 paleotemperature proxy. *International Journal of Plant Science*, 167, 11-18.
902 Royer, D.L., Wilf, P., Janesko, D. A., Kowalski, E. A. & Dilcher, D. L. (2005).
903 Correlations of climate and plant ecology to leaf size and shape: potential
904 proxies for the fossil record. *American Journal of Botany*, 92, 1141-1151.
905 Saarinen, J. (2014). Ecometrics of large herbivorous land mammals in relation to
906 climatic and environmental changes during the Pleistocene. *Department of*
907 *Geosciences and Geography*. Uppsala, Sweden: Uppsala University
908 Dissertation.
909 Saarinen, J. & Karme, A. (2017). Tooth wear and diets of extant and fossil
910 xenarthrans (Mammalia, Xenarthra)–Applying a new mesowear approach.
911 *Palaeogeography, Palaeoclimatology, Palaeoecology*, 476, 42-54.

912 Scholander, P. F. (1955). Evolution of climatic adaptation in homeotherms. *Evolution*,
913 9, 15-26.
914 Sepkoski, J.J. (1978). A kinetic model of Phanerozoic taxonomic diversity I. Analysis
915 of marine orders. *Paleobiology*, 4, 223-251.
916 Shino, Y. & Suzuki, Y. (2011). The ideal hydrodynamic form of the concavo-convex
917 productide brachiopod shell. *Lethaia*, 44, 329-343.
918 Smith, F. A., Lyons, S. K., Ernest, S. K. M., Jones, K. E., Kauffman, D. M., Dayan,
919 T., Marquet, P. A., Brown, J. H. & Haskell, J. P. (2003). Body mass of late
920 Quaternary mammals. *Ecology*, 84, 3403.
921 Stayton, C. T. (2011). Biomechanics on the half shell: functional performance
922 influences patterns of morphological variation in the emydid turtle carapace.
923 *Zoology*, 114, 213–223.
924 Strömberg, C. A. E., Dunn, R. E., Madden, R. H., Kohn, M. J. & Carlini, A. A.
925 (2013). Decoupling the spread of grasslands from the evolution of grazer-type
926 herbivores in South America. *Nature Communications*, 4, 1478.
927 Tapaltsyan, V., Eronen, J.T., Lawing, A.M., Sharir, A., Janis, C., Jernvall, J. & Klein,
928 O.D. (2015). Continuously growing rodent molars result from a predictable
929 quantitative evolutionary change over 50 million years. *Cell reports*, *11*, 673-
930 680.
931 Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier,
932 E. (2007). Let the concept of trait be functional! *Oikos*, 116, 882–892.

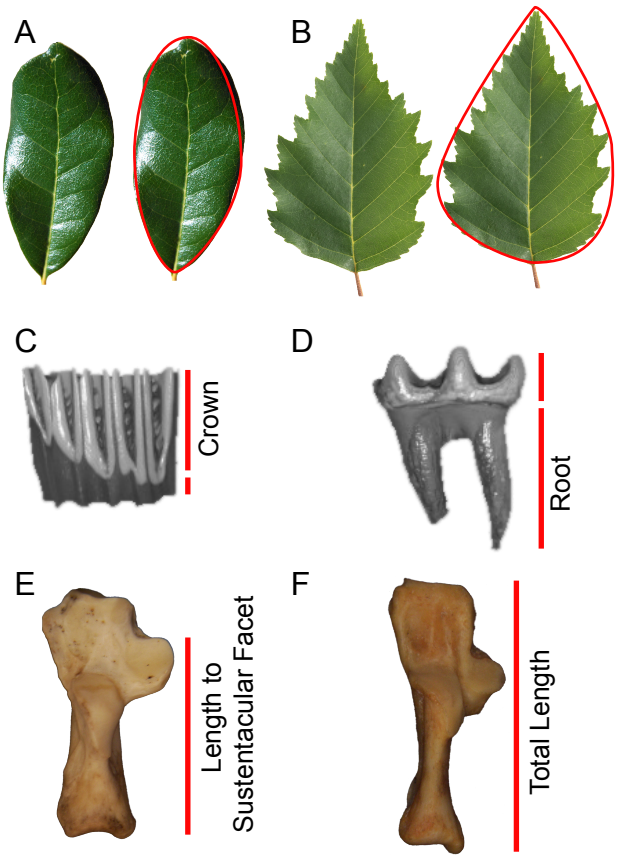
933 Wallin, M. (1991). Ecometric analysis of factors regulating eutrophication effects in
934 coastal waters: a case study of marine fish farms. *Disciplinary Domain of*
935 *Science and Technology*. Uppsala, Sweden: Uppsala University Dissertation.
936 Webster, M. & Hughes, N.C. (1999). Compaction-related deformation in Cambrian
937 olenelloid trilobites and its implications for fossil morphometry. *Journal of*
938 *Paleontology*, 73, 355-371.
939 Wefer, G. & Berger, W. H. (1991). Isotope paleontology: growth and composition of
940 extant calcareous specie. *Marine Geology*, 100, 207-248.
941 Wilf, P. (1997). When are leaves good thermometers? A new case for leaf margin
942 analysis. *Paleobiology*, 23, 373-390.
943 Wilson, G. P., Evans, A. R., Corfe, I. J., Smits, P. D., Fortelius, M. & Jernvall, J.
944 (2012). Adaptive radiation of multituberculate mammals before the extinction
945 of dinosaurs. *Nature*, 483, 457-460.
946 Wing, S. L., Harrington, G. L., Smith, F. A., Bloch, J. I., Boyer, D. M. & Freeman, K.
947 H. (2005). Transient floral change and rapid global warming at the
948 Paleoeocene-Eocene boundary. *Science*, 310, 993-996.
949 Wolfe, J. A. (1979). Temperature parameters of humid to mesic forests of eastern
950 Asia and relation to forests of other regions of the northern hemisphere and
951 Australasia. *Geological Survey Professional Paper*, 1106.
952 —. (1993). A method of obtaining climatic parameters from leaf assemblages.
953 *Bulletin of the US Geological Survey*, 2040.
954 —. (1994). Tertiary climatic changes at middle latitudes of western North America.
955 *Palaeogeography, Palaeoclimatology, Palaeoecology*, 108, 195-205.

956 Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F.,
957 Cavender-Bares, J., *et al.* (2004). The worldwide leaf economics spectrum.
958 *Nature*, 428, 821-827.
959 Wroe, S. (2008). Cranial mechanics compared in extinct marsupial and extant African
960 lions using a finite-element approach. *Journal of Zoology*, 274, 332–339.
961 Yasuhara, M., Tittensor, D. P., Hillebrand, H. & Worm, B. (2015). Combining marine
962 macroecology and palaeoecology in understanding biodiversity: microfossils
963 as a model. *Biological Reviews*, doi: 10.1111/brv.12223
964 Žliobaitė, I., Rinne, J., Tóth, A.B., Mechenich, M., Liu, L., Behrensmeyer, A.K. &
965 Fortelius, M. 2016. Herbivore teeth predict climatic limits in Kenyan
966 ecosystems. *Proceedings of the National Academy of Sciences*, 201609409.

967 **Figure 1.** A. Example of an entire leaf margin from a live oak leaf modified from an
968 image by Jeremy Stovall. The leaf margin is entire when its margin is not punctuated. An
969 easy way to determine if the margin is entire or to quantify the entirety of the margin is to
970 overlay the image of the leaf with an oval, as is depicted in the figure with the red outline.
971 Entire leaf margins are characteristic of hotter environments. B. Example of a non-entire
972 leaf margin from a dwarf birch modified from an image by Fox Valley. Non-entire leaf
973 margins are characteristic of cooler environments. C. Example of hypsodonty in a rodent
974 modified by Tapaltsyan *et al.* 2015. Hypsodonty is measured by the crown to root ratio.
975 Hypsodonts are grazers and are characteristic of arid environments with more gritty
976 vegetation. D. Example of brachydonty in a rodent modified from Tapaltsyan *et al.* 2015.
977 Brachydonts are browsers and are found in more wet environments with mixed
978 vegetation. E. Example of a calcaneum from a raccoon modified from an image on
979 boneid.net. Gear ratio is measured on the calcaneum and is the ratio of the length of the
980 sustentacular facet to the total length of the calcaneum. The gear ratio for the raccoon is
981 typical of a plantigrade animal and is characteristic for animals living in high vegetation
982 cover (e.g., dense woodlands). F. Example of a calcaneum from a panther modified from
983 an image on boneid.net. The gear ratio in the panther is typical of a digitigrade animal
984 and is characteristic of animals that are adapted to run through open habitats.
985
986 **Figure 2.** A. Ecometric space for community-level mammalian body mass showing the
987 maximum likelihood estimation of mean annual temperature (MAT) for each bin. B.
988 Likelihood function for MAT based on mammal body mass at the point in Central
989 America highlighted by the black circles in C-E. C. MAT estimated from mammalian

990 body mass. D. Observed MAT. E. anomaly between estimated and observed MAT. F.
991 Ecometric space for snake body size showing the maximum likelihood estimate for MAT
992 in each bin. G. Likelihood function for MAT at the same point in Central America based
993 on snake size. H. MAT estimated from combined mammalian body mass and snake size.
994 I. Observed MAT. E. Anomaly between combined estimate and observed MAT.
995
996

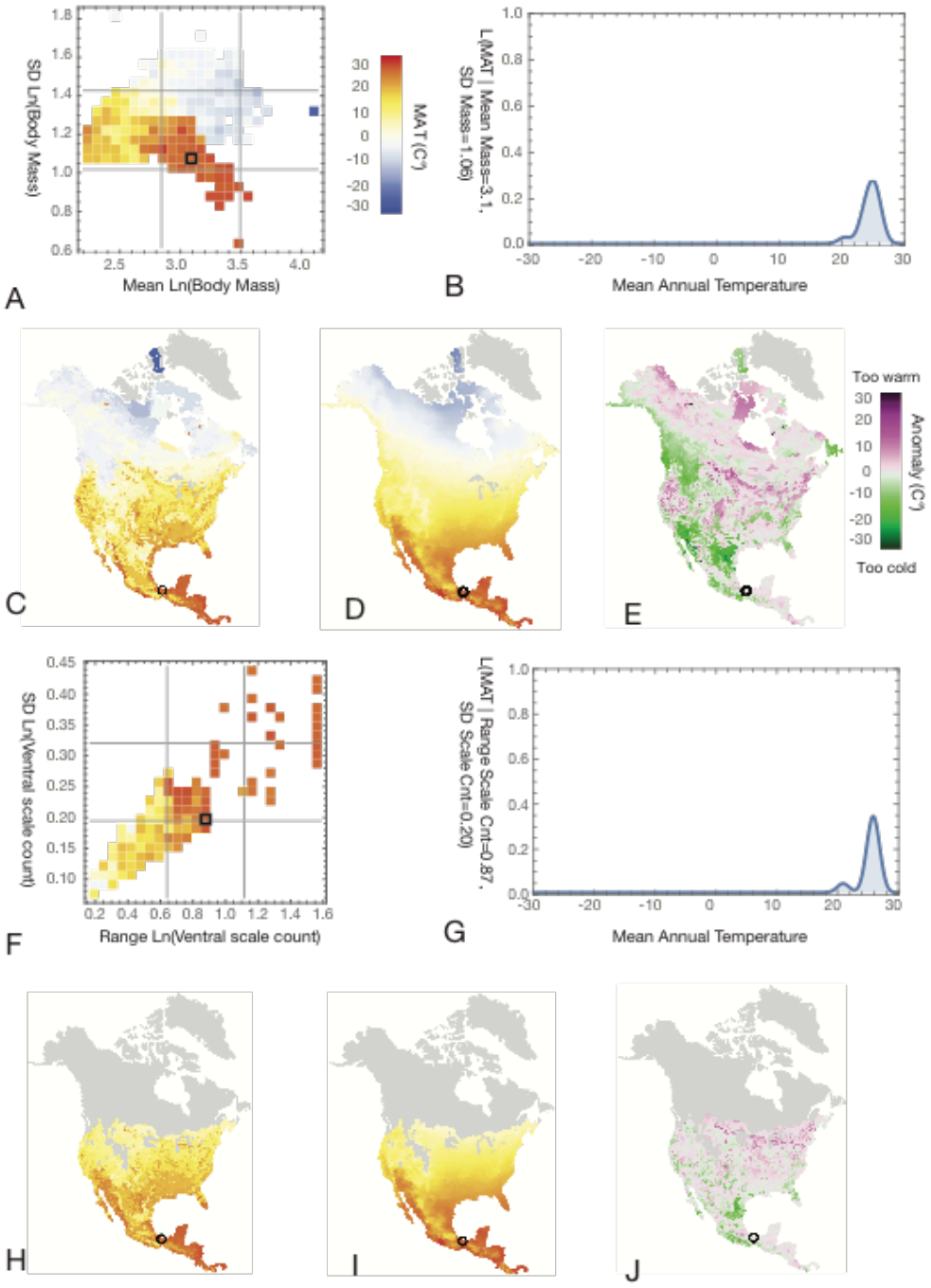
997 **Figure 1**



998

999

1000 **Figure 2**



1001
1002

1003

APPENDIX I. ECOMETRICS WORKFLOW AND R CODE

This section demonstrates an ecometric modeling workflow using the R Statistical Programing Language. To demonstrate these models in R, we will work with spatial data. There are special functions in two packages, *raster* and *sp*, that allow for relatively quick processing of spatial information (Bivand *et al.* 2015, Hijmans 2015, Pebesma and Bivand 2005). We will use climate data from the worldclim database (Hijmans *et al.* 2005) and we will use trait data body mass and hypsodonty, from the PanTHERIA database (Jones *et al.* 2009) and from Eronen *et al.* (2010b). The code below in the gray boxes can be typed directly into an R console or can be entered into an R script file. The blue color indicates that the word is a function. The green color indicates that the text is a character string and the bright blue indicates that the text is recognized as a number. To start the analysis, load the two required libraries. If they are not installed yet on your computer, install them with the function install.packages().

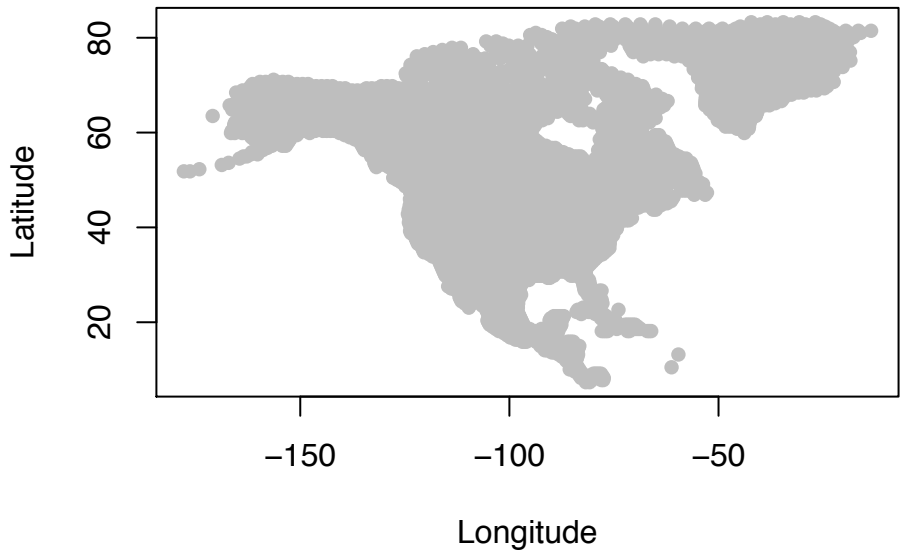
```
library(raster)
## Loading required package: sp
library(sp)
```

Load the sampling locations and look at the first six rows of data with the functions read.csv() and head(). The first function read.csv() is a wrapper for another function called read.table(), which can be used in place of read.csv(), if the data are in tab delimited format. Use the help() function to see the documentation associated with each function.

```
points <- read.csv("data/SamplingPoints.csv")
head(points)
##   GLOBALID Longitude Latitude
## 1   103148  -42.1727 83.26264
## 2   103149  -38.3442 83.26264
## 3   103150  -34.5156 83.26264
## 4   103151  -30.6871 83.26264
## 5   103152  -26.8586 83.26264
## 6   103235  -79.4690 82.81348
```

Plot the sampling locations with the function plot() to visualize the geographic distribution of the sampling locations. In this example, we use 50 km equidistant points sampled across North America. These are the same points used in Polly (2010).

```
plot(points[,2:3], col = "gray", pch = 16)
```



Download raster climate data from the worldclim database using the getData() function from the package raster that we loaded with the library() function (Hijmans *et al.* 2005, Hijmans 2015). In this example, we download the 10 minute resolution, but if you would like to try a higher resolution data set, then change the argument named res to 2.5 or 0.5. Extract the temperature and precipitation for each sampling location using the extract() function.

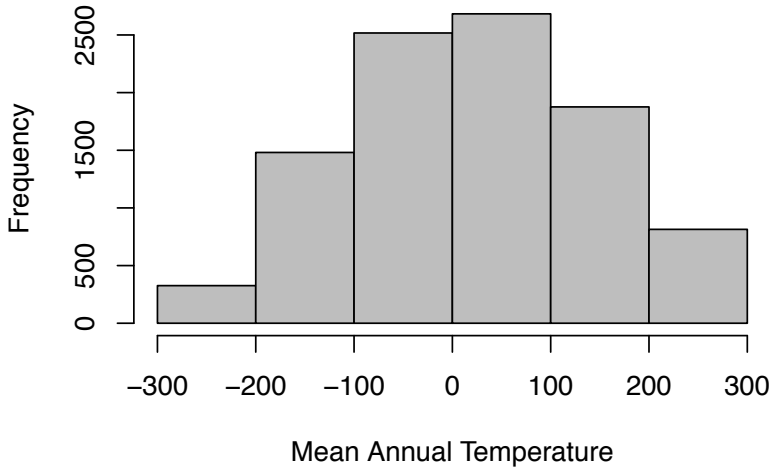
```
bioclim <- getData("worldclim", download = T, path = "data", var = "bio", res = 10)
```

Extract the temperature for each sampling location.

```
temperature <- extract(bioclim[[1]], points[,2:3])
```

Calculate the temperature range for all the sampling localities to make a plot of the temperature. We add one to the range to make the range equal to index values that we can use to subset the color function. The R language starts the subset of data at an index value of 1. Calculate the color value associated with each temperature value and the temperature values associated with even breaks to assign legend values.

```
temp_range <- 1 + max(temperature, na.rm = T) - min(temperature, na.rm = T)
colfunc_temp <- colorRampPalette(c("darkblue", "blue", "gray", "yellow",
"red"))(temp_range)[1 + temperature - min(temperature, na.rm = T)]
h <- hist(temperature, main = "", xlab = "Mean Annual Temperature", col = "gray",
breaks = 5)
```



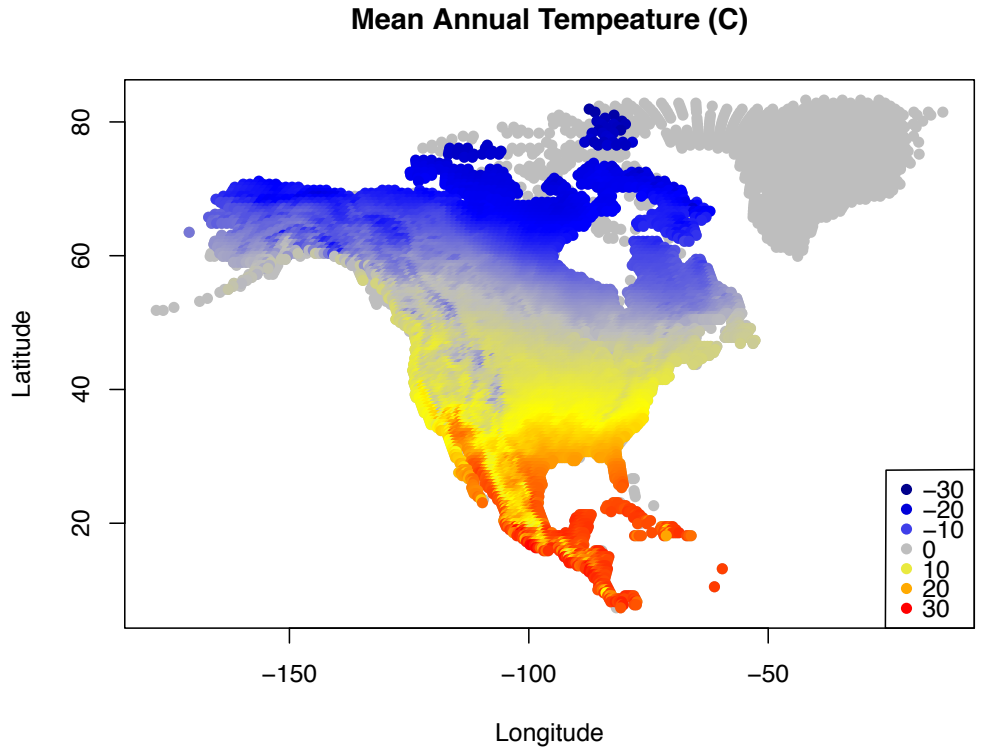
1064

1065 **plot**(points[,2:3], col = colfunc_temp, pch = 16, main = "Mean Annual Tempeature (C)")

1066 **legend**("bottomright", legend = h$breaks/10, pch = 16, col =

1067 **colorRampPalette**(**c**("darkblue", "blue", "gray", "yellow", "red"))(**length**(h$breaks)))

1068



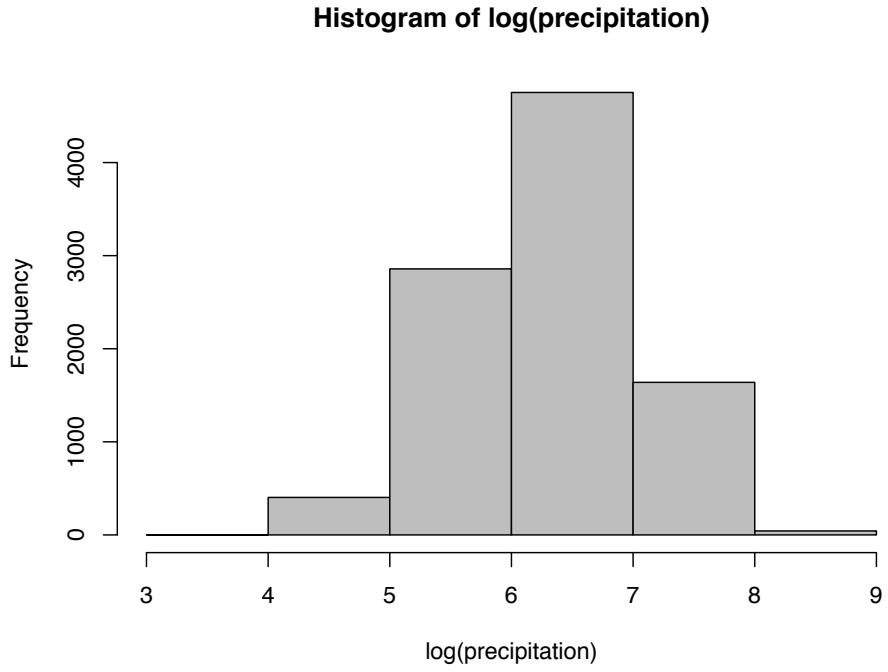
1069

Extract the precipitation for each sampling locality.

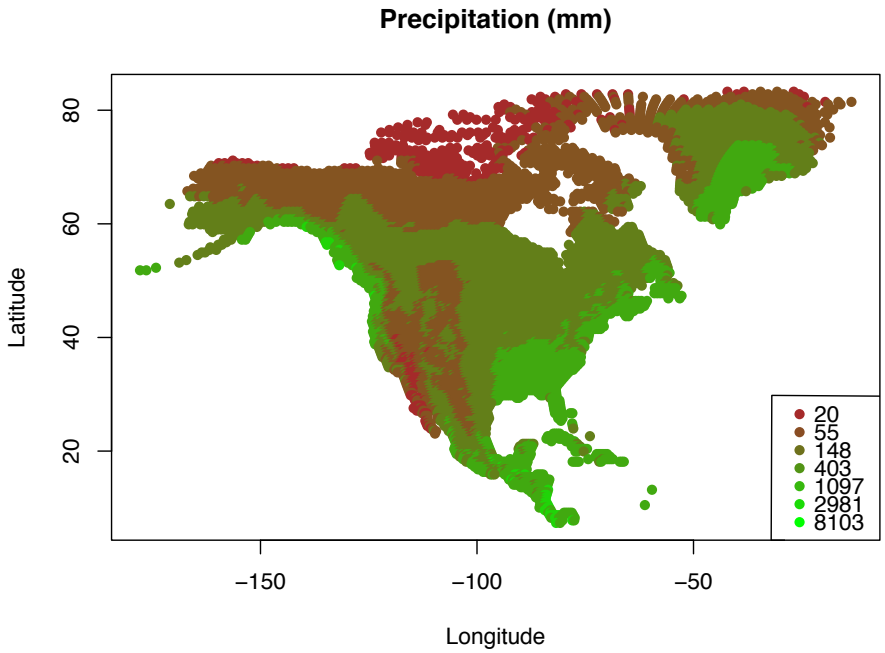
```
precipitation <- extract(bioclim[[12]], points[,2:3])
```

Calculate the precipitation range for all the sampling localities to make a plot of the precipitation. Also, calculate color value associated with each precipitation value and the precipitation values associated with even breaks to assign legend values.

```
precip_range <- 1 + max(log(precipitation), na.rm = T) - min(log(precipitation), na.rm = T)
colfunc_pr <- colorRampPalette(c("brown", "green"))(precip_range)[1 + log(precipitation) - min(log(precipitation), na.rm = T)]
h <- hist(log(precipitation), breaks = 5, col = "gray")
```



```
plot(points[,2:3], col = colfunc_pr, pch = 16, main = "Precipitation (mm)")
legend(-36.25, 60.5, legend = round(exp(h$breaks)), pch = 16, col =
colorRampPalette(c("brown", "green"))(length(h$breaks)))
```



Compile the climate variables into a new data.frame called climate. Remove the variables that are taking up memory with the rm() function if your memory is getting sluggish.

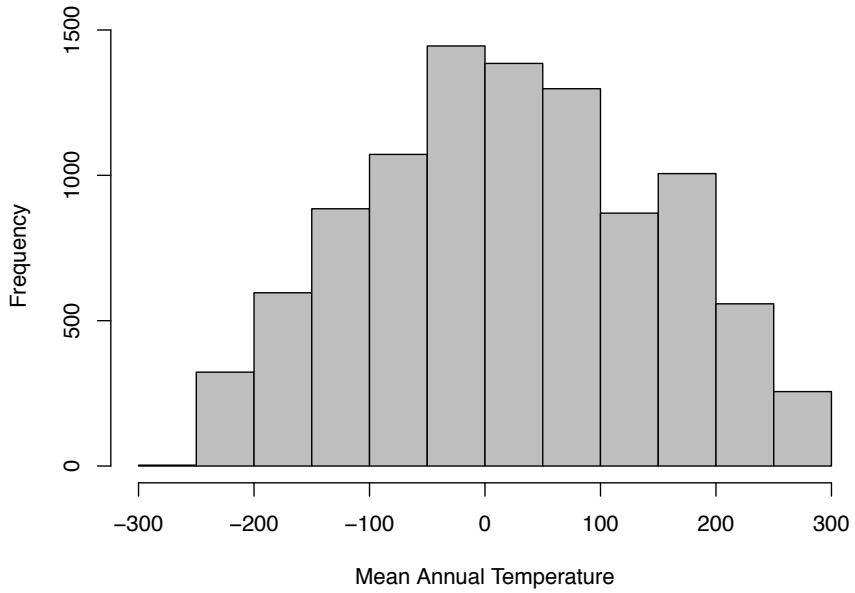
```
climate <- cbind(points, temperature, precipitation)
#rm(bioclim, temperature, precipitation, points)
```

Visually check the climate variables for normality and if they are not mostly normally distributed, transform them for normality.

```
head(climate)
```

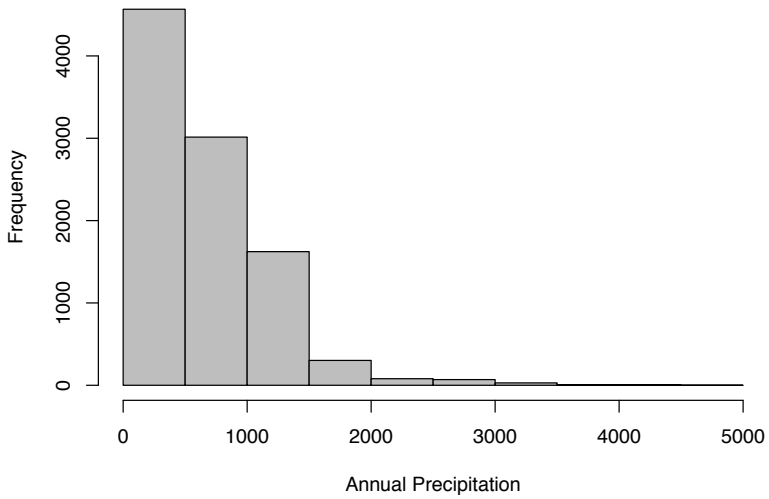
```
##   GLOBALID Longitude Latitude temperature precipitation
## 1   103148  -42.1727 83.26264        -169         139
## 2   103149  -38.3442 83.26264        -170         141
## 3   103150  -34.5156 83.26264        -175         149
## 4   103151  -30.6871 83.26264        -185         166
## 5   103152  -26.8586 83.26264        -180         139
## 6   103235  -79.4690 82.81348        -207          90
```

```
hist(climate[,4], main = "", xlab = "Mean Annual Temperature", col = "gray")
```



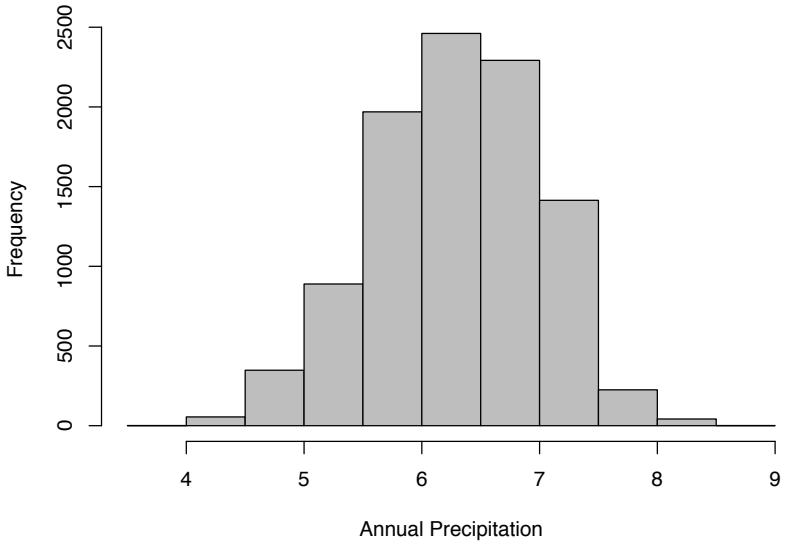
Temperature appears to be mostly normally distributed, so now we check precipitation.

```
hist(climate[,5], main = "", xlab = "Annual Precipitation", col = "gray")
```



Precipitation appears to be log distributed. We log transform this variable to get it closer to normality.

```
climate[,5] <- log(climate[,5])
hist(climate[,5], main = "", xlab = "Annual Precipitation", col = "gray")
```



Next, we read in the trait data from a folder called data. We assign the row names of the new data frame to the names of the taxon within the dataset. We look at the first six rows of the trait data frame with the head() function. The two traits that we use in this example are body mass and hypsodonty index. Body mass is reported in grams and is the mass of any adult reported in the PanTHERIA database (Jones *et al.* 2009) from live or freshly-killed specimens. These include captive, wild, provisioned, or unspecified populations and include male, female, and sex unspecified individuals. The mean for each species is reported for each species. The second trait that we use is and index for hypsodonty from Eronen *et al.* (2010b).

```
traits <- read.csv("data/NAmammalTraits.csv")
rownames(traits) <- traits$TaxonName
head(traits)
##                              TaxonName BodyMass hypsodonty_index
## Didelphis virginiana     Didelphis virginiana 3.387760              1
## Aplodontia rufa              Aplodontia rufa 2.906448              3
## Sciurus carolinensis     Sciurus carolinensis 2.736715              1
## Sciurus griseus              Sciurus griseus 2.847480              1
## Sciurus niger                Sciurus niger       NA              1
## Tamiasciurus douglasii Tamiasciurus douglasii 2.352183              1
```

Now we read in shapefiles containing polygons that represent the geographic ranges for all of the species of interest. These specific shape files were obtained from IUCN Redlist

using their spatial data download option (www.iucnredlist.org). If you are dealing with large shapefiles, then this step will take a reasonable amount of processing time.

```
geography <-
shapefile("data/TERRESTRIAL_MAMMALS/TERRESTRIAL_MAMMALS.shp")
```

Next we create a list of species at each sampling locality by first turning the sampling points into spatial points with the function SpatialPoints(). We assign the coordinate reference system of our spatial points to a proj4string to match the coordinate reference system of the spatial polygons representing the geographic ranges. We then create a list with the function over(). If you are dealing with large shapefiles, keep in mind that the over() function will take a reasonable amount of time to process.

```
sp <- SpatialPoints(climate[,2:3], proj4string = CRS(proj4string(geography)))
o <- over(sp, geography, returnList = T)
```

The sample size at each site is calculated by determining the length of the vector returned for each site. The ecometric for body mass and hypsodonty index are summarized for the community level distribution. Here, we summarize with the mean.

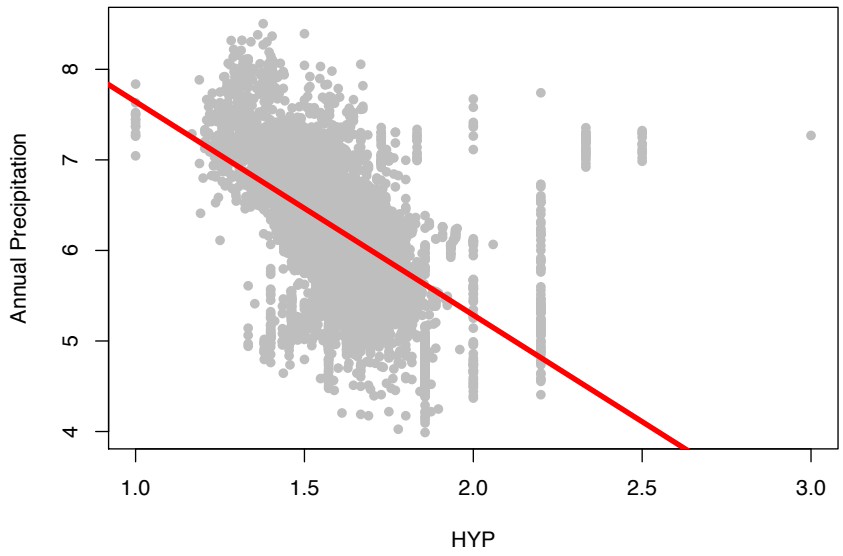
```
richness <- unlist(lapply(o, function(x) length(traits[x$binomial,"hypsodonty_index"])))
ecometric_bodymass <- unlist(lapply(o, function(x)
mean(traits[x$binomial,"BodyMass"], na.rm = T)))
ecometric_hypsodonty <- unlist(lapply(o, function(x)
mean(traits[x$binomial,"hypsodonty_index"], na.rm = T)))
```

FIRST APPROXIMATION WITH TRANSFER FUNCTION

Now we create a model describing the relationship between traits and climate. First we consider the relationship between hypsodonty and precipitation. We build a simple linear model to describe the variation in precipitation due to the variation in hypsodonty using the function lm(). We only use sites that we have data for more than five species. We look at a summary of the model using the function summary(). Both the intercept and the coefficient (here the coefficient represents the slope of the linear relationship) are not zero ($p < 0.001$). The amount of explained variation (R^2) is 30%. We then make a scatterplot of those variables to look at the general spread of data and add the linear model with the function abline().

```
model_hyp <- lm(climate[richness > 5,5] ~ ecometric_hypsodonty[richness > 5])
summary(model_hyp)
## Call:
## lm(formula = climate[richness > 5, 5] ~ ecometric_hypsodonty[richness >  5])
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9984 -0.3348 -0.0163  0.3821  4.3388
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       9.99886    0.06160   162.3   <2e-16
## ecometric_hypsodonty[richness > 5] -2.35602    0.03869   -60.9   <2e-16
## (Intercept)                       ***
## ecometric_hypsodonty[richness > 5] ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 0.5834 on 8651 degrees of freedom
##   (15 observations deleted due to missingness)
## Multiple R-squared:  0.3001, Adjusted R-squared:    0.3
## F-statistic:  3709 on 1 and 8651 DF,  p-value: < 2.2e-16
plot(ecometric_hypsodonty[richness > 5], climate[richness > 5,5], xlab = "HYP", ylab =
"Annual Precipitation", pch = 16, col = "gray")
abline(model_hyp, col = "red", lwd = 4)
```

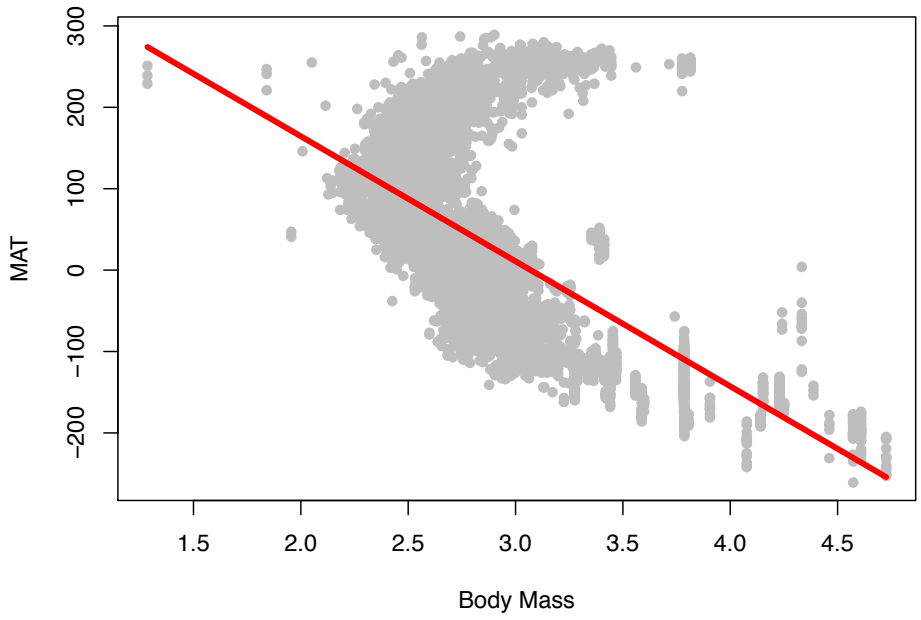


From this model, we can see there is some predictive power in this transfer function, but the linear model does not capture the relationship well. In the next section we will show how to estimate annual precipitation from hypsodonty with a maximum likelihood approach that better captures the relationship between annual precipitation and hypsodonty.

Now we create a model describing the relationship between the body mass and temperature. We build a linear model to describe the variation in body mass due to the variation in temperature using the function lm(). We look at a summary of the model using the function summary(). Both the intercept and all the coefficients are significantly different from zero ($p < 0.001$). The amount of explained variation (R2) is approximately 34%. We then make a scatterplot of those variables to look at the general spread of data and add the model with the function curve().

```
model_mass <- lm(climate[richness > 5,4] ~ ecometric_bodymass[richness > 5])
summary(model_mass)
## Call:
## lm(formula = climate[richness > 5, 4] ~ ecometric_bodymass[richness >  5])
## Residuals:
##     Min      1Q  Median      3Q     Max
## -170.72  -65.22  -24.73   43.71  375.19
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       471.512      6.571   71.76   <2e-16 ***
## ecometric_bodymass[richness > 5] -153.559      2.304  -66.64   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## Residual standard error: 92.6 on 8651 degrees of freedom
##   (15 observations deleted due to missingness)
```

```
## Multiple R-squared:  0.3392, Adjusted R-squared:  0.3391
## F-statistic:  4440 on 1 and 8651 DF,  p-value: < 2.2e-16
plot(ecometric_bodymass[richness > 5], climate[richness > 5,4], ylab = "MAT", xlab =
"Body Mass", pch = 16, col = "gray")
curve(model_mass$coefficients[1] + model_mass$coefficients[2] * x, col = "red", lwd =
4, add = T)
```



From this model, we can see that, again, there is some predictive power in this transfer function, but the linear model does not capture the relationship well. In the next section we will show how to estimate mean annual temperature from body mass with a maximum likelihood approach that better captures the relationship between the two.

The coefficients that were estimated in both of these models can be used to estimate paleotemperature and precipitation. Confidence limits can also be calculated given the input dataset. It is important to note that the size of the confidence limits will vary with

climate. For example, between 5 C and 28 C, there is a stronger relationship with body size than below or above those temperatures. Hypsodonty has high variability throughout the precipitation range present in North America; however, there is a central tendency about the average relationship between precipitation and hypsodonty index that is useful in reconstructing paleoprecipitation with confidence limits.

MAXIMUM LIKELIHOOD ESTIMATION

Although transfer functions, while easy to apply and adequate for first approximations, assume a fairly simple relationship one-to-one relationship between environment and trait means. Combining different traits that have functional relationships with the same environmental factor is also awkward with conventional regression-based transfer functions, especially if the traits are fundamentally different in kind or scale (e.g., body mass measured in kg and humerus shape measured in Procrustes units).

An alternative strategy is to estimate the likelihood of environmental parameters given the distribution of traits in a community (Lawing *et al.* 2012; Polly and Head 2015). This approach, like many likelihood or Bayesian methods, requires far fewer assumptions about the statistical distributions of variables and it allows otherwise incommensurable data to be combined into the same estimate.

To begin, we need to create another variable at the community level, namely the standard deviation, to use in the maximum likelihood estimate of temperature.

```
sd_ecometric_bodymass <- unlist(lapply(o, function(x)
sd(traits[x$binomial,"BodyMass"], na.rm = T)))
```

We create bins using the body mass variable and extract the break points for each bin.

```
#bin the community level trait distribution into 25X25
#first take the range of each
mtemp <- range(ecometric_bodymass, na.rm = T)
sdtemp <- range(sd_ecometric_bodymass, na.rm = T)
#get the break points for the mean and sd
mbrks <- seq(mtemp[1], mtemp[2], diff(mtemp)/25)
sdbrks <- seq(sdtemp[1], sdtemp[2], diff(sdtemp)/25)
#assign bin codes for each
mbc <- .bincode(ecometric_bodymass, breaks = mbrks)
sdbc <- .bincode(sd_ecometric_bodymass, breaks = sdbrks)
```

We calculate the temperature for each bin.

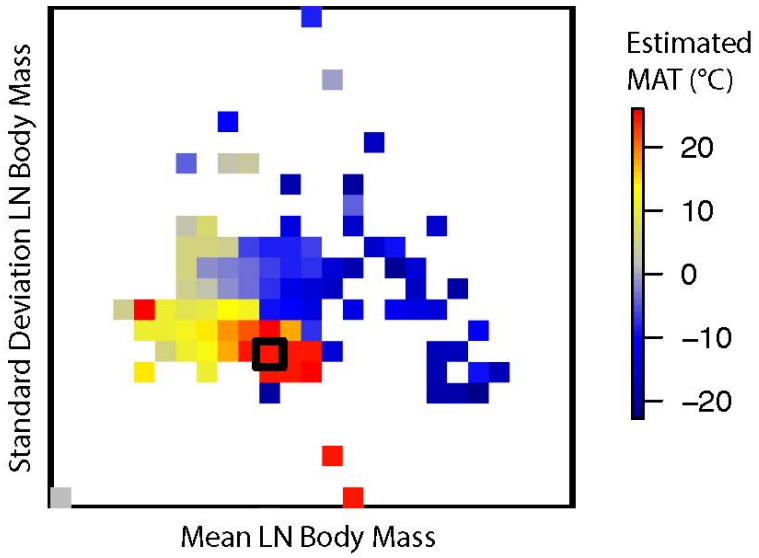
```
#calculate the data for the raster
obj <- array(NA,dim = c(25,25))
for(i in 1:25){
  for(j in 1:25){
    dat <- round(temperature[which(mbc==i & sdbc==j)]/10)
    obj[26 - j,i] <- mean(dat, na.rm = T)
  }
}
```

Next, we create a raster to store the body mass and temperature data for bins.

```
#make a raster
r <- raster(extent(0,25,0,25), resolution = 1)
#set the values to the obj
r <- setValues(r,obj)
```

Plot the raster and highlight the bin that we will use to extract data to show an example of that maximum likelihood estimate.

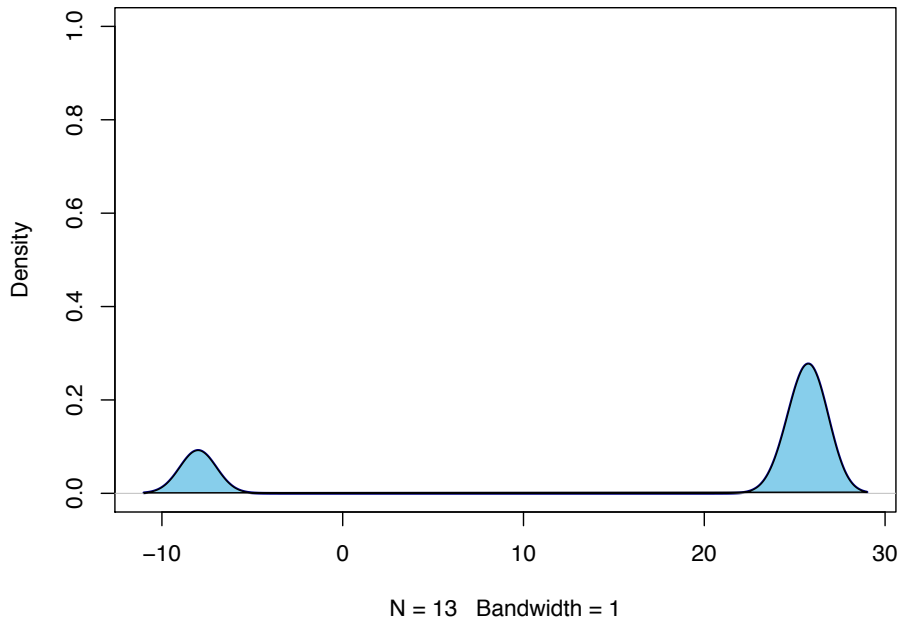
```
#make an empty plot
plot(1:25, 1:25, type = "n", xlim = c(1,25), ylim = c(1,25),
    xaxs = "i", yaxs = "i", asp = 1, axes = F, xlab ="", ylab="")
#add the rectangle/box
rect(0, 1, 25, 25, lwd = 3)
#add the raster data
plot(r, col = colorRampPalette(c("darkblue", "blue",
   "grey","yellow", "red"))(round(maxValue(r) -
   minValue(r))), add = T)
#this is mean = 3.1, 12, and sd = 1.08, 10
rect(11, 9, 12, 10, lwd = 4)
```



The colors in this raster plot show the Mean Annual Temperature maximum likelihood estimate given the associated mean and standard deviation of each bin.

We extract the data for the highlighted bin and plot the kernel density with a Gaussian kernel. This shows the distribution of the likelihood surface.

```
#grab all the data for that box
dat <- round(temperature[which(mbc==12 & sdbc==10)]/10)
#plot the kernel density with gaussian kernel, bandwidth = 1
mod <- density(dat, bw = 1)
plot(mod, ylim = c(0,1), col = "darkblue", lwd = 2)
polygon(mod$x, mod$y, col = "skyblue")
```



This likelihood surface shows a bimodal distribution of the most likely temperature. Although it is bimodal, it is much more likely that the temperature falls on the warm end of the spectrum, as opposed to the cold end.

Next, we calculate the maximum likelihood for all bins.

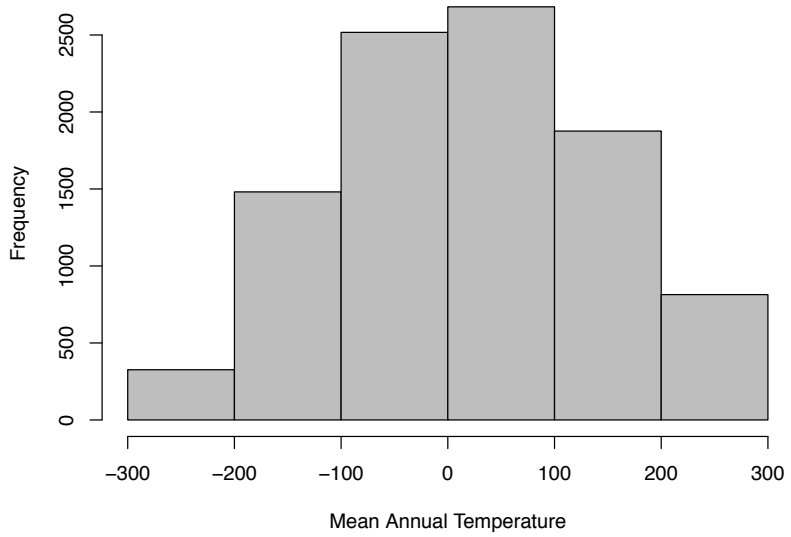
```
modmax <- array(NA, dim = length(points[,1]))
mod <- list()
for(i in 1:length(points[,1])){
  if(!(is.na(mbc[i]) | is.na(sdbc[i]))){
    dat <- round(temperature[which(mbc==mbc[i] & sdbc==sdbc[i])]/10)
    mod[[i]] <- density(dat, bw = 1)
    modmax[i] <- mod[[i]]$x[which.max(mod[[i]]$y)]
}}
modmax <- round(modmax*10)
```

We only use bins with more than the number of species specified as the cutoff. Here we use seven. This means that there needs to be at least seven species recorded at each location to be included in the estimate.

```
cutoff <- 7
```

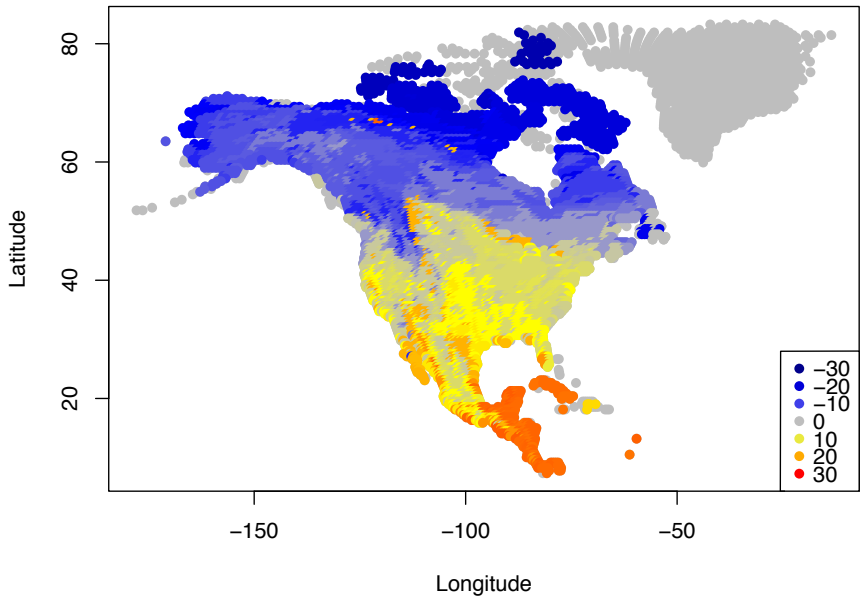
To plot the maximum likelihood temperature estimate from the ecometric values, we create a color palette for the temperature estimates. In addition, we save the histogram with five break points to a variable to use in plotting.

```
colfunc_eco <- colorRampPalette(c("darkblue", "blue", "gray", "yellow",
"red"))(temp_range)[1 + modmax - min(modmax, na.rm = T)]
h <- hist(temperature, main = "", xlab = "Mean Annual Temperature", col = "gray",
breaks = 5)
```



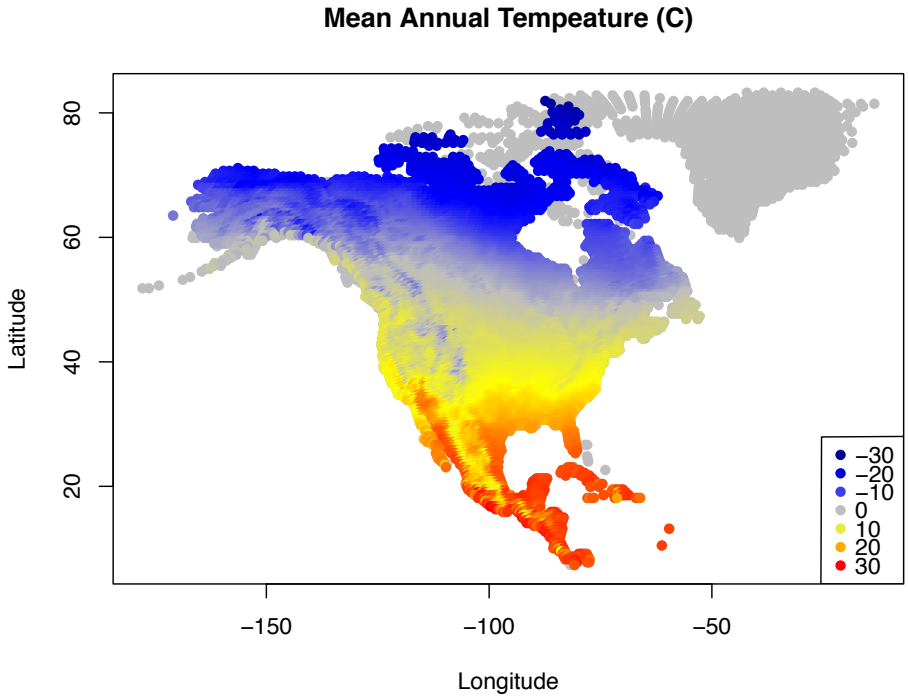
We plot the maximum likelihood temperature estimate from body mass.

```
plot(points[,2:3], col = "gray", pch = 16)
points(points[richness > cutoff, 2:3], col = colfunc_eco[richness > cutoff], pch = 16)
legend(-31.5, 61, legend = h$breaks/10, pch = 16, col = colorRampPalette(c("darkblue",
"blue", "gray", "yellow", "red"))(length(h$breaks)))
```



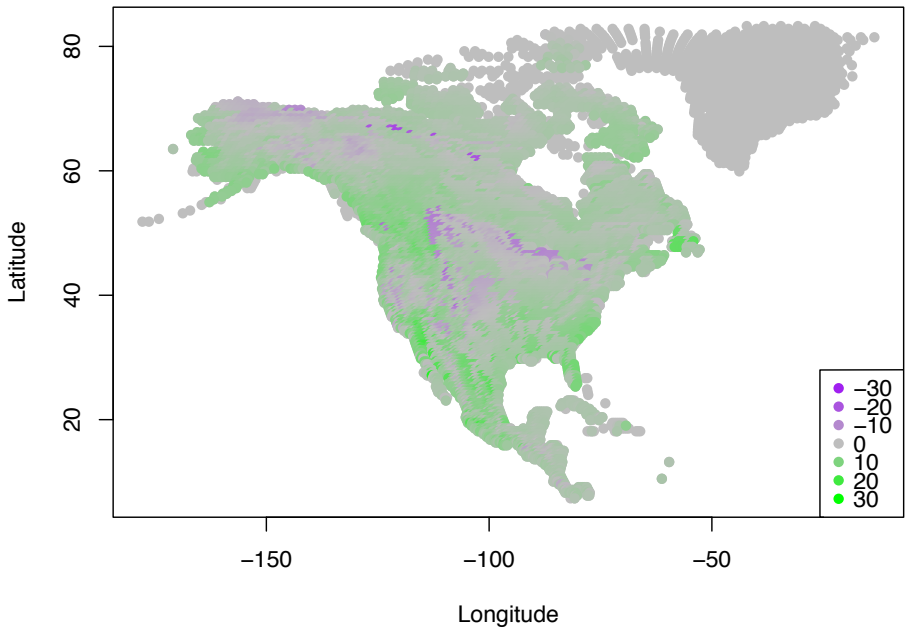
Next we plot the actual temperature to compare with the estimated temperature.

```
plot(points[,2:3], col = "gray", pch = 16, main = "Mean Annual Tempeature (C)")
points(points[richness > cutoff,2:3], col = colfunc_temp[richness > cutoff], pch = 16)
legend(-31.5, 61, legend = h$breaks/10, pch = 16, col = colorRampPalette(c("darkblue",
"blue", "gray", "yellow", "red"))(length(h$breaks)))
```



We plot the anomaly to visualize the difference between the estimated and actual Mean annual Temperature.

```
plot(points[,2:3], col = "gray", pch = 16)
anom <- temperature - modmax
colfunc_anom <- colorRampPalette(c("purple", "grey", "green"))(max(anom, na.rm = T) - min(anom, na.rm = T))[1 + anom - min(anom, na.rm = T)]
points(points[richness > cutoff, 2:3], col = colfunc_anom[richness > cutoff], pch = 16)
legend(-31.5, 61, legend = h$breaks/10, pch = 16, col = colorRampPalette(c("purple", "grey", "green"))(length(h$breaks)))
```



The anomaly between the estimated Mean Annual Temperature and the actual Mean Annual Temperature shows that most of the temperature estimates are less that one degree C.

APPENDIX I. REFERENCES

Bivand, R. S., Pebesma, E. & Gomez-Rubio, V. (2013). *Applied spatial data analysis with R,* Second edition. New York, NY: Springer.

Eronen, J., Puolamaki, K., Liu, L., Lintulaakso, K., Damuth, J., Janis, C. & Fortelius, M. (2010). Precipitation and large herbivorous mammals I: estimates from present-day communities. *Evolutionary Ecology Research*, 12, 217-233.

Hijmans, R. J. (2015). raster: Geographic Data Analysis and Modeling. *R package version 2.3-40.*

Hijmans, R. J., Cameron, J. L., Parra, P. G., Jones, P. G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965-1978.

Jones, K. E., Bielby, J., Cardillo, M., Fritz, S. A., O'Dell, J., Orme, C. D. L., Safi, K., *et al.* (2009). PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, 90, 2648-2648.

Lawing, A. M., Head, J. J. & Polly, P. D. (2012). The ecology of morphology: the ecometrics of locomotion and macroenvironment in North American snakes. In J. Louys (Eds). *Paleontology in Ecology and Conservation* (pp. 117-146). Berlin: Springer Berlin Heidelberg.

Pebesma, E. J. & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, 5, 2.

Polly, P. D. & Head, J. J. (2015). Measuring Earth-life transitions: ecometric analysis of functional traits from late Cenozoic vertebrates. In P. D. Polly, J. J. Head, & D. L. Fox (Eds.), *Earth-Life Transitions: Paleobiology in the Context of Earth System Evolution* (pp. 21-46). The Paleontological Society Papers 21. New Haven, CT: Yale Press.